

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



547 447

(43) Internationales Veröffentlichungsdatum
10. September 2004 (10.09.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/076617 A2

- (51) Internationale Patentklassifikation⁷: C12N
- (21) Internationales Aktenzeichen: PCT/EP2004/000771
- (22) Internationales Anmeldedatum:
29. Januar 2004 (29.01.2004)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
103 08 836.9 27. Februar 2003 (27.02.2003) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von
US): BASF PLANT SCIENCE GMBH [DE/DE]; 67056
Ludwigshafen (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): RENZ, Andreas
[DE/DE]; Heinrich-von-Kleist-Strasse 6, 67117 Limburg-
erhof (DE). HEINZ, Ernst [DE/DE]; Püttkampsweg 13,
22609 Hamburg (DE). ABBADI, Amine [DE/DE]; Lüb-
bersmeyer Weg 26, 22549 Hamburg (DE). DOMERGUE,
Frederic [FR/DE]; Bahrenfelder Steindamm 98, 22761
Hamburg (DE). ZANK, Thorsten [DE/DE]; Secken-
heimer Str. 4-6, 68165 Mannheim (DE).
- (74) Anwalt: PRESSLER, Uwe; c/o BASF Aktienge-
sellschaft, 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,
CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare regionale Schutzrechtsart): ARIPO (BW,
GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM,
ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu ver-
öffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Ab-
kürzungen wird auf die Erklärungen ("Guidance Notes on Co-
des and Abbreviations") am Anfang jeder regulären Ausgabe der
PCT-Gazette verwiesen.

(54) Title: METHOD FOR THE PRODUCTION OF POLYUNSATURATED FATTY ACIDS

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG MEHRFACH UNGESÄTTIGTER FETTSÄUREN

(57) Abstract: The invention relates to a method for producing polyunsaturated fatty acids in an organism by introducing nucleic acids coding for polypeptides with acyl-CoA:lysophospholipid-acyl transferase activity into the organism. Advantageously, said nucleic acid sequences can optionally be expressed in the transgenic organism along with additional nucleic acid sequences coding for polypeptides of fatty acid metabolism or lipid metabolism biosynthesis. The invention also relates to the nucleic acid sequences, nucleic acid structures containing said nucleic acid sequences, vectors containing the inventive nucleic acid sequences, and/or the invention further relates to oils, lipids, and/or fatty acids that are produced according to the inventive method, and the use thereof.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren in einem Organismus, indem Nukleinsäuren in den Organismus eingebracht werden, die für Polypeptide mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codieren. Vorteilhaft können diese Nukleinsäuresequenzen gegebenenfalls zusammen mit weiteren Nukleinsäuresequenzen, die für Polypeptide der Biosynthese des Fettsäure- oder Lipidstoffwechsels codieren, in dem transgenen Organismus exprimiert werden. Die Erfindung betrifft weiterhin die Nukleinsäuresequenzen, Nukleinsäurekonstrukte enthaltend die erfindungsgemäßen Nukleinsäuresequenzen, Vektoren enthaltend die Nukleinsäuresequenzen und/oder die Nukleinsäurekonstrukte sowie transgene Organismen enthaltend die vorgenannten Nukleinsäuresequenzen, Nukleinsäurekonstrukte und/oder Vektoren. Ein weiterer Teil der Erfindung betrifft Öle, Lipide und/oder Fettsäuren hergestellt nach dem erfindungsgemäßen Verfahren und deren Verwendung.

WO 2004/076617 A2

Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren

Beschreibung

- Die vorliegende Erfindung betrifft ein Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren in einem Organismus indem Nukleinsäuren in den Organismus eingebracht werden, die für Polypeptide mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codieren. Vorteilhaft können diese Nukleinsäuresequenzen gegebenenfalls zusammen mit weiteren Nukleinsäuresequenzen, die für Polypeptide der Biosynthese des Fettsäure- oder Lipidstoffwechsels codieren, in dem transgenen Organismus exprimiert werden.
- 10 Die Erfindung betrifft weiterhin die Nukleinsäuresequenzen, Nukleinsäurekonstrukte enthaltend die erfindungsgemäßen Nukleinsäuresequenzen, Vektoren enthaltend die Nukleinsäuresequenzen und/oder die Nukleinsäurekonstrukte sowie transgene Organismen, die vorgenannten Nukleinsäuresequenzen, Nukleinsäurekonstrukte und/oder Vektoren.
- 15 Ein weiterer Teil der Erfindung betrifft Öle, Lipide und/oder Fettsäuren hergestellt nach dem erfindungsgemäßen Verfahren und deren Verwendung.

- Fettsäuren und Triglyceride haben eine Vielzahl von Anwendungen in der Lebensmittelindustrie, der Tierernährung, der Kosmetik und im Pharmabereich. Je nachdem ob es sich um freie gesättigte oder ungesättigte Fettsäuren oder um Triglyceride mit einem erhöhten Gehalt an gesättigten oder ungesättigten Fettsäuren handelt, sind sie für die unterschiedlichsten Anwendungen geeignet, so werden beispielsweise mehrfach ungesättigte Fettsäuren Babynahrung zur Erhöhung des Nährwertes zugesetzt. Mehrfach ungesättigte ω -3-Fettsäuren und ω -6-Fettsäuren stellen dabei einen wichtigen Bestandteil der tierischen und menschlichen Nahrung dar. Aufgrund der heute üblichen Zusammensetzung der menschlichen Nahrung ist ein Zusatz von mehrfach ungesättigten ω -3-Fettsäuren, die bevorzugt in Fischölen vorkommen, zur Nahrung besonders wichtig. So werden beispielsweise mehrfach ungesättigte Fettsäuren wie Docosahexaensäure (= DHA, C22:6^{A4,7,10,13,16,19}) oder Eicosapentaensäure (= EPA, C20:5^{A5,8,11,14,17}) Babynahrung zur Erhöhung des Nährwertes zugesetzt. Der ungesättigten Fettsäure DHA wird dabei ein positiver Effekt auf die Entwicklung des Gehirns zugeschrieben.

Im folgenden werden mehrfach ungesättigte Fettsäuren als PUFA, PUFAs, LCPUFA oder LCPUFAs bezeichnet (poly unsaturated fatty acids, **PUFA**, long chain poly unsaturated fatty acids, **LCPUFA**).

35. Hauptsächlich werden die verschiedenen Fettsäuren und Triglyceride aus Mikroorganismen wie Mortierella oder Schizochytrium oder aus Öl-produzierenden Pflanzen wie Soja, Raps, Algen wie Crypthecodinium oder Phaeodactylum und weiteren gewonnen, wobei sie in der Regel in Form ihrer Triacylglyceride (= Triglyceride = Triglycerole) anfallen. Sie können aber auch aus Tieren wie z.B. Fischen gewonnen werden. Die freien

5. Fettsäuren werden vorteilhaft durch Verseifung hergestellt. Höhere mehrfach ungesättigte Fettsäuren wie DHA, EPA, Arachidonsäure (= ARA, C20:4^{Δ5,8,11,14}), Dihomo-γ-linolensäure (C20:3^{Δ8,11,14}) oder Docosapentaensäure (DPA, C22:5^{Δ7,10,13,16,19}) lassen sich nicht aus Ölfruchtpflanzen wie Raps, Soja, Sonnenblume, Färberdistel oder anderen isolieren. Übliche natürliche Quellen für diese Fettsäuren sind Fische wie Hering, Lachs, Sardine, Goldbarsch, Aal, Karpfen, Forelle, Heilbutt, Makrele, Zander oder Thunfisch oder Algen.

10 Je nach Anwendungszweck sind Öle mit gesättigten oder ungesättigten Fettsäuren bevorzugt, so sind z.B. in der humanen Ernährung Lipide mit ungesättigten Fettsäuren speziell mehrfach ungesättigten Fettsäuren bevorzugt. Den mehrfach ungesättigten ω-3-Fettsäuren wird dabei ein positiver Effekt auf den Cholesterinspiegel im Blut und damit auf die Möglichkeit der Prävention einer Herzerkrankung zugeschrieben. Durch Zugabe dieser ω-3-Fettsäuren zu Nahrung kann das Risiko einer Herzerkrankung, eines Schlaganfalls oder von Bluthochdruck deutlich verringert werden. Auch entzündliche 15 speziell chronisch entzündliche Prozesse im Rahmen immunologischer Erkrankungen wie rheumatoider Arthritis lassen sich durch ω-3-Fettsäuren positiv beeinflussen. Sie werden deshalb Lebensmitteln speziell diätischen Lebensmitteln zugegeben oder finden in Medikamenten Anwendung. ω-6-Fettsäuren wie Arachidonsäure haben bei diesen rheumatischen Erkrankungen aufgrund unserer üblichen Nahrungsmittelzusammensetzung eher einen negativen Effekt auf diese Krankheiten. 20

ω-3- und ω-6-Fettsäuren sind Vorläufer von Gewebshormonen, den sogenannten Eicosanoiden wie den Prostaglandinen, die sich von der Dihomo-γ-linolensäure, der Arachidonsäure und der Eicosapentaensäure ableiten, den Thromoxanen und Leukotrienen, die sich von der Arachidonsäure und der Eicosapentaensäure ableiten. Eicosanoide (sog. PG₂-Serie), die aus ω-6-Fettsäuren gebildet werden fördern in der Regel 25 Entzündungsreaktionen, während Eicosanoide (sog. PG₃-Serie) aus ω-3-Fettsäuren geringe oder keine entzündungsfördernde Wirkung haben.

Aufgrund ihrer positiven Eigenschaften hat es in der Vergangenheit nicht an Ansätzen gefehlt, Gene, die an der Synthese von Fettsäuren bzw. Triglyceriden beteiligt sind, für 30 die Herstellung von Ölen in verschiedenen Organismen mit geändertem Gehalt an ungesättigten Fettsäuren verfügbar zu machen. So wird in WO 91/13972 und seinem US-Äquivalent eine Δ-9-Desaturase beschrieben. In WO 93/11245 wird eine Δ-15-Desaturase in WO 94/11516 wird eine Δ-12-Desaturase beansprucht. Weitere Desaturasen werden beispielsweise in EP-A-0 550 162, WO 94/18337, WO 97/30582, 35 WO 97/21340, WO 95/18222, EP-A-0 794 250, Stukey et al., J. Biol. Chem., 265, 1990: 20144-20149, Wada et al., Nature 347, 1990: 200-203 oder Huang et al., Lipids 34, 1999: 649-659 beschrieben. Die biochemische Charakterisierung der verschiedenen Desaturasen ist jedoch bisher nur unzureichend erfolgt, da die Enzyme als membrangebundene Proteine nur sehr schwer zu isolieren und zu charakterisieren sind 40 (McKeon et al., Methods in Enzymol. 71, 1981: 12141-12147, Wang et al., Plant Physiol. Biochem., 26, 1988: 777-792). In der Regel erfolgt die Charakterisierung membrangebundener Desaturasen durch Einbringung in einen geeigneten Organismus, der

anschließend auf Enzymaktivität mittels Edukt- und Produktanalyse untersucht wird. Δ -6-Desaturasen werden in WO 93/06712, US 5,614,393, US5614393, WO 96/21022, WO00/21557 und WO 99/27111 beschrieben und auch die Anwendung zur Produktion in transgenen Organismen beschrieben wie in WO98/46763 WO98/46764, WO9846765. Dabei wird auch die Expression verschiedener Desaturasen wie in WO99/64616 oder WO98/46776 und Bildung polyungesättigter Fettsäuren beschrieben und beansprucht. Bzgl. der Effektivität der Expression von Desaturasen und ihren Einfluss auf die Bildung polyungesättigter Fettsäuren ist anzumerken, dass durch Expression einer einzelnen Desaturase wie bisher beschrieben lediglich geringe Gehalte an ungesättigten Fettsäuren/Lipiden wie z.B. γ -Linolensäure und Stearidonsäure erreicht wurden. Weiterhin wurde in der Regel ein Gemisch aus ω -3- und ω -6-Fettsäuren erhalten.

Besonders geeignete Mikroorganismen zur Herstellung von PUFAs sind Mikroorganismen wie Thraustochytrien oder Schizochytrien-Stämme, Algen wie *Phaeodactylum tricornutum* oder *Cryptocodinium*-Arten, Ciliaten, wie *Stylonychia* oder *Colpidium*, Pilze, wie *Mortierella*, *Entomophthora* oder *Mucor*. Durch Stammselektion ist eine Anzahl von Mutantenstämmen der entsprechenden Mikroorganismen entwickelt worden, die eine Reihe wünschenswerter Verbindungen, einschließlich PUFAs, produzieren. Die Mutation und Selektion von Stämmen mit verbesserter Produktion eines bestimmten Moleküls wie den mehrfach ungesättigten Fettsäuren ist jedoch ein zeitraubendes und schwieriges Verfahren. Deshalb werden, wenn immer möglich wie oben beschrieben gentechnologische Verfahren bevorzugt. Mit Hilfe der vorgenannten Mikroorganismen lassen sich jedoch nur begrenzte Mengen der gewünschten mehrfach ungesättigten Fettsäuren wie DPA, EPA oder ARA herstellen. Wobei diese in der Regel je nach verwendeten Mikroorganismus als Fettsäuregemische aus beispielsweise EPA, DPA und DHA anfallen.

Alternativ kann die Produktion von Feinchemikalien im großen Maßstab vorteilhaft über die Produktion in Pflanzen durchgeführt werden, die so entwickelt werden, dass sie die vorstehend genannten PUFAs herstellen. Besonders gut für diesen Zweck geeignete Pflanzen sind Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten wie Raps, Canola, Lein, Soja, Sonnenblumen, Borretsch und Nachtkerze. Aber auch andere Nutzpflanzen, die Öle oder Lipide und Fettsäuren enthalten, sind gut geeignet, wie in der eingehenden Beschreibung dieser Erfindung erwähnt. Mittels herkömmlicher Züchtung ist eine Reihe von Mutantepflanzen entwickelt worden, die ein Spektrum an wünschenswerten Lipiden und Fettsäuren, Cofaktoren und Enzymen produzieren. Die Selektion neuer Pflanzensorten mit verbesserter Produktion eines bestimmten Moleküls ist jedoch ein zeitaufwändiges und schwieriges Verfahren oder sogar unmöglich, wenn die Verbindung in der entsprechenden Pflanze nicht natürlich vorkommt, wie im Fall von mehrfach ungesättigten C_{18} -, C_{20} -Fettsäuren und C_{22} -Fettsäuren und solchen mit längeren Kohlenstoffketten.

Aufgrund der positiven Eigenschaften ungesättigter Fettsäuren hat es in der Vergangenheit nicht an Ansätzen gefehlt, diese Gene, die an der Synthese von Fettsäuren

bzw. Triglyceriden beteiligt sind, für die Herstellung von Ölen in verschiedenen Pflanzen mit einem geänderten Gehalt an mehrfach ungesättigten Fettsäuren verfügbar zu machen. Bisher konnten jedoch länger-kettige mehrfach ungesättigte C₂₀- und/oder C₂₂-Fettsäuren wie EPA oder ARA nicht in Pflanzen hergestellt werden.

- 5 Aber auch in anderen Organismen wie Mikroorganismen wie Algen oder Pilzen führten die gentechnischen Veränderungen des Fettsäurestoffwechselweges über das Einbringen und die Expression beispielsweise von Desaturasen nur zu relativ geringen Steigerungen der Produktivität in diesen Organismen. Ein Grund hierfür mag in dem sehr komplexen Fettsäurestoffwechsel liegen. So ist der Einbau von mehrfach ungesättigten Fettsäuren in Membranlipide und/oder in Triacylglyceride und deren Ab- und Umbau sehr komplex und bis heute biochemisch und speziell genetisch noch nicht vollständig aufgeklärt und verstanden.

- Die Biosynthese von LCPUFAs und der Einbau von LCPUFAs in Membranen oder Triacylglyceride erfolgt über verschiedene Stoffwechselwege (Abbadi et al. (2001) European Journal of Lipid Science & Technology 103:106-113). In Bakterien wie *Vibrio* und Mikroalgen wie *Schizochytrium* wird Malonyl-CoA über eine LCPUFA-produzierende Polyketidsynthase zu LCPUFAs umgesetzt (Metz et al. (2001) Science 293: 290-293; WO 00/42195; WO 98/27203; WO 98/55625). In Mikroalgen wie *Phaeodactylum* und Moosen wie *Physcomitrella* werden ungesättigte Fettsäuren wie Linolensäure oder Linolensäure in Form ihrer Acyl-CoAs in mehreren Desaturierungs- und Elongationsschritten zu LCPUFAs umgesetzt (Zank et al. (2000) Biochemical Society Transactions 28: 654-658). Bei Säugetieren beinhaltet die Biosynthese von DHA zusätzlich zu Desaturierungs- und Elongationsschritten eine Kettenverkürzung über β -Oxidation.

- 25 LCPUFAs liegen in Mikroorganismen und niederen Pflanzen entweder ausschließlich in Form von Membranlipiden vor, wie bei *Physcomitrella* und *Phaeodactylum* oder sie sind in Membranlipiden und Triacylglyceriden vorhanden, wie bei *Schizochytrium* und *Mortierella*. Der Einbau von LCPUFAs in Lipide und Öle wird durch verschiedene - Acyltransferasen und Transacylasen katalysiert. Diese Enzyme sind bereits bekannt für den Einbau von gesättigten und ungesättigten Fettsäuren [Slabas (2001) J. Plant Physiology 158: 505-513; Frentzen (1998) Fett/Lipid 100: 161-166; Cases et al. (1998) Proc. Nat. Acad. Sci. USA 95: 13018-13023]. Bei den Acyltransferasen handelt es sich um Enzyme des sogenannten Kennedy-Pathways, die an der cytoplasmatischen Seite des Membransystems des Endoplasmatischen Reticulums, nachfolgend als 'ER' bezeichnet, lokalisiert sind. Experimentell können Membranen des ER als sogenannte 'mikrosomale Fraktionen' aus verschiedenen Organismen isoliert werden (Knutzon et al. (1995) Plant Physiology 109: 999-1006; Mishra & Kamisaka (2001) Biochemistry 355: 315-322; US 5968791). Diese ER-gebundenen Acyltransferasen in der mikrosomalen Fraktion verwenden Acyl-CoA als aktivierte Form der Fettsäuren. Glycerin-3-phosphat Acyltransferase, im folgenden GPAT genannt, katalysiert den Einbau von Acylgruppen an der sn-1 Position von Glycerin-3-phosphat. 1-Acylglycerin-3-phosphat Acyltransferase (E.C. 2.3.1.51), auch Lysophosphatidsäure Acyltransferase, im folgenden LPAAT

genannt, katalysiert den Einbau von Acylgruppen an der sn-2 Position von Lysophosphatidsäure, nachfolgend als LPA abgekürzt. Nach Dephosphorylierung von Phosphatidsäure durch Phosphatidsäure Phosphatase katalysiert Diacylglycerin Acyltransferase, im folgenden DAGAT genannt, den Einbau von Acylgruppen an der sn-3 Position von Diacylglycerins. Neben diesen Kennedy Pathway Enzymen sind weitere Enzyme am Einbau von Fettsäuren in Triacylglyceride beteiligt, die Acylgruppen aus Membranlipiden in Triacylglyceride einbauen können. Phospholipid Diacylglycerin Acyltransferase, nachfolgend PDAT genannt, und Lysophosphatidylcholin Acyltransferase, nachfolgend LPCAT genannt.

- 10 Die enzymatische Aktivität einer LPCAT wurde erstmals in Ratten beschrieben [Land (1960) *Journal of Biological Chemistry* 235: 2233-2237]. In Pflanzen existiert eine plastidäre Isoform der LPCAT [Akermoun et al. (2000) *Biochemical Society Transactions* 28: 713-715] sowie eine ER gebundene Isoform [Tumaney und Rajasekharan (1999) *Biochimica et Biophysica Acta* 1439: 47-56; Fraser und Stobart, *Biochemical Society Transactions* (2000) 28: 715-7718]. LPCAT ist in Tieren wie auch in Pflanzen an der Biosynthese und der Transacylierung von mehrfach ungesättigten Fettsäuren beteiligt [Stymne und Stobart (1984) *Biochem. J.* 223: 305-314; Stymne und Stobart (1987) in 'The Biochemistry of Plants: a Comprehensive Treatise', Vol. 9 (Stumpf, P.K. ed.) pp. 175-214, Academic Press, New York]. Eine wichtige Funktion der LPCAT oder allgemeiner gesagt einer Acyl-CoA:Lysophospholipid Acyltransferase, nachfolgend LPLAT genannt, bei der ATP-unabhängigen Synthese von Acyl-CoA aus Phospholipiden wurde von Yamashita et al. (2001; *Journal of Biological Chemistry* 276: 26745-26752) beschrieben.

- 25 Trotz vieler biochemischer Daten konnten bisher keine Gene kodierend für LPCAT identifiziert werden. Gene anderer verschiedener pflanzlicher Acyltransferasen konnten isoliert werden und werden in WO 00/18889 (Novel Plant Acyltransferases) beschrieben.

- 30 Höhere Pflanzen enthalten mehrfach ungesättigte Fettsäuren wie Linolsäure (C18:2) und Linolensäure (C18:3). Arachidonsäure (ARA), Eicosapentaensäure (EPA) und Docosahexaensäure (DHA) kommen wie oben beschrieben im Samenöl höherer Pflanzen gar nicht oder nur in Spuren vor (E. Ucciani: *Nouveau Dictionnaire des Huiles Végétales. Technique & Documentation – Lavoisier*, 1995. ISBN: 2-7430-0009-0). Es ist vorteilhaft, in höheren Pflanzen, bevorzugt in Ölsaaten wie Raps, Lein, Sonnenblume und Soja, LCPUFAs herzustellen, da auf diese Weise große Mengen qualitativ hochwertiger LCPUFAs für die Lebensmittelindustrie, die Tierernährung und für pharmazeutische Zwecke kostengünstig gewonnen werden können. Hierzu werden vorteilhaft über gentechnische Methoden Gene kodierend für Enzyme der Biosynthese von LCPUFAs in Ölsaaten eingeführt und exprimiert werden. Dies sind beispielsweise Gene kodierend für Δ -6-Desaturase, Δ -6-Elongase, Δ -5-Desaturase, Δ -5-Elongase und Δ -4-Desaturase. Diese Gene können vorteilhaft aus Mikroorganismen, Tieren und niederen Pflanzen isoliert werden, die LCPUFAs herstellen und in den Membranen oder Triacylglyceriden einbauen. So konnten bereits Δ -6-Desaturase-Gene aus dem Moos

Physcomitrella patens und Δ -6-Elongase-Gene aus *P. patens* und dem Nematoden *C. elegans* isoliert.

5 Erste transgene Pflanzen, die Gene kodierend für Enzyme der LCPUFA-Biosynthese enthalten und exprimieren und LCPUFAs produzieren wurden beispielsweise in DE 102 19 203 (Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren in Pflanzen) erstmals beschrieben. Diese Pflanzen produzieren allerdings LCPUFAs in Mengen, die für eine Aufarbeitung der in den Pflanzen enthaltenen Öle noch weiter optimiert werden müssen.

10 Um eine Anreicherung der Nahrung und des Futters mit diesen mehrfach ungesättigten Fettsäuren zu ermöglichen, besteht daher ein großer Bedarf an einem einfachen, kostengünstigen Verfahren zur Herstellung dieser mehrfach ungesättigten Fettsäuren speziell in eukaryontischen Systemen.

15 Es bestand daher die Aufgabe ein Verfahren zur Herstellung von mehrfach ungesättigten Fettsäuren in einem eukaryontischen Organismus zu entwickeln. Diese Aufgabe wurde durch das erfindungsgemäße Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren in einem Organismus, dadurch gekennzeichnet, dass das Verfahren - folgende Schritte umfasst:

- 20 a) Einbringen mindestens einer Nukleinsäuresequenz in den Organismus mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Sequenz, die für ein Polypeptid mit einer Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codiert; oder
- 25 b) Einbringen mindestens einer Nukleinsäuresequenz in den Organismus, die sich als Ergebnis des degenerierten genetischen Codes von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 enthaltenden codierenden Sequenz ableiten lässt, oder
- 30 c) Einbringen mindestens eines Derivates der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Nukleinsäuresequenz in den Organismus, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 dargestellten Aminosäuresequenz codieren und mindestens 40 % Homologie auf Aminosäureebene mit SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 aufweisen und eine äquivalente Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität aufweisen, und
- d) kultivieren und ernten des Organismus.

35 Vorteilhaft enthalten die im erfindungsgemäßen Verfahren hergestellten mehrfach ungesättigten Fettsäuren mindestens zwei vorteilhaft drei Doppelbindungen. Besonders vorteilhaft enthalten die Fettsäuren vier oder fünf Doppelbindungen. Im Verfahren hergestellte Fettsäuren haben vorteilhaft 16-, 18-, 20- oder 22 C-Atome in der Fettsäure-

kette. Diese hergestellten Fettsäuren können als einziges Produkt im Verfahren hergestellt werden oder in einem Fettsäuregemisch vorliegen.

5 Bei den im erfindungsgemäßen Verfahren verwendeten Nukleinsäuresequenzen handelt es sich um isolierte Nukleinsäuresequenzen, die für Polypeptide mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codieren.

Die im Verfahren hergestellten mehrfach ungesättigten Fettsäuren sind vorteilhaft in Membranlipiden und/oder Triacylglyceriden gebunden, können aber auch als freie Fettsäuren oder aber gebunden in Form anderer Fettsäureester in den Organismen vor-
kommen. Dabei können sie wie gesagt als "Reinprodukte" oder aber vorteilhaft in Form
10 von Mischungen verschiedener Fettsäuren oder Mischungen unterschiedlicher Glyceride vorliegen. Dabei lassen sich die in den Triacylglyceriden gebundenen verschiedenen Fettsäuren von kurzkettigen Fettsäuren mit 4 bis 6 C-Atomen, mittelkettigen Fettsäuren mit 8 bis 12 C-Atomen oder langkettigen Fettsäuren mit 14 bis 24 C-Atomen ableiten, bevorzugt sind die langkettigen Fettsäuren besonders bevorzugt sind die langkettigen
15 Fettsäuren LCPUFAs von C₁₈-, C₂₀- und/oder C₂₂-Fettsäuren.

Im erfindungsgemäßen Verfahren werden vorteilhaft Fettsäureester mit mehrfach ungesättigten C₁₆-, C₁₈-, C₂₀- und/oder C₂₂-Fettsäuremolekülen mit mindestens zwei Doppelbindungen im Fettsäureester hergestellt. Bevorzugt enthalten diese Fettsäuremoleküle drei, vier oder fünf Doppelbindungen und führen vorteilhaft zur Synthese von
20 Hexadecadiensäure (C16:2^{Δ^{9,12}}), γ-Linolensäure (= GLA, C18:3^{Δ^{6,9,12}}), Stearidonsäure (= SDA, C18:4^{Δ^{6,9,12,15}}), Dihomo-γ-Linolensäure (= DGLA, 20:3^{Δ^{8,11,14}}), Eicosatetraensäure (= ETA, C20:4^{Δ^{5,8,11,14}}), Arachidonsäure (ARA), Eicosapentaensäure (EPA) oder deren Mischungen, bevorzugt EPA und/oder ARA.

Die Fettsäureester mit mehrfach ungesättigten C₁₆-, C₁₈-, C₂₀- und/oder C₂₂-
25 Fettsäuremolekülen können aus den Organismen, die für die Herstellung der Fettsäureester verwendet wurden, in Form eines Öls oder Lipids beispielsweise in Form von Verbindungen wie Sphingolipide, Phosphoglyceride, Lipide, Glycolipide wie Glycosphingolipid, Phospholipide wie Phosphatidylethanolamin, Phosphatidylcholin, Phosphatidylserin, Phosphatidylglycerol, Phosphatidylinositol oder Diphosphatidylglycerol, Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Fettsäureester wie die AcetylCoenzymA-Ester, die die mehrfach ungesättigten Fettsäuren mit mindestens zwei bevorzugt drei Doppelbindungen enthalten, isoliert werden. Neben diesen Estern sind die mehrfach ungesättigten Fettsäuren auch als freie Fettsäuren oder ge-
30 bunden in anderen Verbindungen in den Organismen vorteilhaft den Pflanzen enthalten. In der Regel liegen die verschiedenen vorgenannten Verbindungen (Fettsäureester und frei Fettsäuren) in den Organismen in einer ungefähren Verteilung von 80 bis 90 Gew.-% Triglyceride, 2 bis 5 Gew.-% Diglyceride, 5 bis 10 Gew.-% Monoglyceride, 1 bis 5 Gew.-% freie Fettsäuren, 2 bis 8 Gew.-% Phospholipide vor, wobei sich die Summe der verschiedenen Verbindungen zu 100 Gew.-% ergänzt.
35

40 Im erfindungsgemäßen Verfahren werden die hergestellten LCPUFAs mit einem Gehalt von mindestens 3 Gew.-%, vorteilhaft von mindestens 5 Gew.-%, bevorzugt von

mindestens 8 Gew.-%, besonders bevorzugt von mindestens 10 Gew.-%, ganz besonders bevorzugt von mindestens 15 Gew.-% bezogen auf die gesamten Fettsäuren in der transgenen Organismen vorteilhaft in einer transgenen Pflanze hergestellt. Da im erfindungsgemäßen Verfahren von den Ausgangsverbindungen Hexadecadiensäure (C16:2), Linolsäure (C18:2) bzw. Linolensäure (C18:3) mehrere Reaktionsschritte durchlaufen werden, fallen die Endprodukte des Verfahrens wie beispielsweise Arachidonsäure (ARA) oder Eicosapentaensäure (EPA) nicht als absolute Reinprodukte an, es sind immer auch geringe Spuren der Vorstufen im Endprodukt enthalten. Sind in dem Ausgangsorganismus bzw. in der Ausgangspflanze beispielsweise sowohl Linolsäure als auch Linolensäure vorhanden, so liegen die Endprodukte wie ARA und EPA als Mischungen vor. Die Vorstufen sollten vorteilhaft nicht mehr als 20 Gew.-%, bevorzugt nicht mehr als 15 Gew.-%, besonders bevorzugt nicht als 10 Gew.-%, ganz besonders bevorzugt nicht mehr als 5 Gew.-% bezogen auf die Menge des jeweiligen Endprodukts betragen. Vorteilhaft werden in einer transgenen Pflanze als Endprodukte nur ARA oder nur EPA im erfindungsgemäßen Verfahren gebunden oder als freie Säuren hergestellt. Werden beide Verbindungen (ARA + EPA) gleichzeitig hergestellt, werden sie vorteilhaft in einem Verhältnis von mindestens 1:2 (EPA:ARA), vorteilhaft von mindestens 1:3, bevorzugt von 1:4, besonders bevorzugt von 1:5 hergestellt.

Durch die erfindungsgemäßen Nukleinsäuresequenzen kann eine Steigerung der Ausbeute an mehrfach ungesättigten Fettsäuren von mindestens 50 %, vorteilhaft von mindestens 80 %, besonders vorteilhaft von mindestens 100 %, ganz besonders vorteilhaft von mindestens 150 % gegenüber den nicht transgenen Ausgangsorganismus beim Vergleich in der GC-Analyse erreicht werden.

Auch chemisch reine mehrfach ungesättigte Fettsäuren oder Fettsäurezusammensetzungen sind nach den vorbeschriebenen Verfahren darstellbar. Dazu werden die Fettsäuren oder die Fettsäurezusammensetzungen aus dem Organismus wie den Mikroorganismen oder den Pflanzen oder dem Kulturmedium, in dem oder auf dem die Organismen angezogen wurden, oder aus dem Organismus und dem Kulturmedium in bekannter Weise beispielsweise über Extraktion, Destillation, Kristallisation, Chromatographie oder Kombinationen dieser Methoden isoliert. Diese chemisch reinen Fettsäuren oder Fettsäurezusammensetzungen sind für Anwendungen im Bereich der Lebensmittelindustrie, der Kosmetikindustrie und besonders der Pharmaindustrie vorteilhaft.

Als Organismus für die Herstellung im erfindungsgemäßen Verfahren kommen prinzipiell alle Organismen wie Pilze wie *Mortierella* oder *Traustochytrium*, Hefen wie *Saccharomyces* oder *Schizosaccharomyces*, Moose wie *Physcomitrella* oder *Ceratodon*, nicht-humane Tiere wie *Caenorhabditis*, Algen wie *Cryptocodinium* oder *Phaeodactylum* oder Pflanzen wie zweikeimblättrige oder einkeimblättrige Pflanzen in Frage. Vorteilhaft werden Organismen im erfindungsgemäßen Verfahren verwendet, die zu den Öl-produzierenden Organismen gehören, das heißt die für die Herstellung von Ölen verwendet werden, wie Pilze wie *Mortierella* oder *Traustochytrium*, Algen wie *Cryptocodinium*, *Phaeodactylum* oder Pflanzen, insbesondere Pflanzen bevorzugt Ölfrucht-

pflanzen, die große Mengen an Lipidverbindungen enthalten, wie Erdnuss, Raps, Canola, Sonnenblume, Safflor (Färberdistel), Mohn, Senf, Hanf, Rizinus, Olive, Sesam, Calendula, Punica, Nachtkerze, Königskerze, Distel, Wildrosen, Haselnuss, Mandel, Macadamia, Avocado, Lorbeer, Kürbis, Lein, Soja, Pistazien, Borretsch, Bäume (Ölpalme, Kokosnuss oder Walnuss) oder Feldfrüchte, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa oder Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten sowie ausdauernde Gräser und Futterfeldfrüchte. Bevorzugte erfindungsgemäße Pflanzen sind Ölfuchtpflanzen, wie Erdnuss, Raps, Canola, Sonnenblume, Safflor (Färberdistel), Mohn, Senf, Hanf, Rhizinus, Olive, Calendula, Punica, Nachtkerze, Kürbis, Lein, Soja, Borretsch, Bäume (Ölpalme, Kokosnuss). Besonders bevorzugt sind C18:2- und/oder C18:3-Fettsäure reiche Pflanzen wie Sonnenblume, Färberdistel, Tabak, Königskerze, Sesam, Baumwolle, Kürbis, Mohn, Nachtkerze, Walnuss, Lein, Hanf, Distel oder Färberdistel. Ganz besonders bevorzugt sind Pflanzen wie Färberdistel, Sonnenblume, Mohn, Nachtkerze, Walnuss, Lein oder Hanf.

Für das erfindungsgemäße beschriebene Verfahren ist es vorteilhaft in den Organismus zusätzlich zu den unter Verfahrensschritt (a) bis (c) eingebrachten Nukleinsäuren zusätzlich weitere Nukleinsäuren einzubringen, die für Enzyme des Fettsäure- oder Lipidstoffwechsels codieren.

Im Prinzip können alle Gene des Fettsäure- oder Lipidstoffwechsels vorteilhaft in Kombination mit der erfinderischen Acyl-CoA:Lysophospholipid-Acyltransferase im Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren verwendet werden vorteilhaft werden Gene des Fettsäure- oder Lipidstoffwechsels ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) in Kombination mit der Acyl-CoA:Lysophospholipid-Acyltransferase verwendet. Besonders bevorzugt werden Gene ausgewählt aus der Gruppe der Δ -4-Desaturasen, Δ -5-Desaturasen, Δ -6-Desaturasen, Δ -8-Desaturasen, Δ -9-Desaturasen, Δ -12-Desaturasen, Δ -5-Elongasen, Δ -6-Elongasen oder Δ -9-Elongasen in Kombination mit der Acyl-CoA:Lysophospholipid-Acyltransferase im erfindungsgemäßen Verfahren verwendet.

Durch die enzymatische Aktivität der im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren, die für Polypeptide mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codieren, vorteilhaft in Kombination mit Nukleinsäuresequenzen, die für Polypeptide des Fettsäure- oder Lipidstoffwechsels wie Δ -4-, Δ -5-, Δ -6-, Δ -8-Desaturase- oder Δ -5-, Δ -6- oder Δ -9-Elongaseaktivität codieren, können unterschiedlichste mehrfach ungesättigte Fettsäuren im erfindungsgemäßen Verfahren hergestellt werden. Je nach Auswahl der für das erfindungsgemäße Verfahren verwendeten Organismen

wie den vorteilhaften Pflanze lassen sich Mischungen der verschiedenen mehrfach ungesättigten Fettsäure oder einzelne mehrfach ungesättigte Fettsäuren wie EPA oder ARA in freier oder gebundener Form herstellen. Je nachdem welche Fettsäurezusammensetzung in der Ausgangspflanze vorherrscht (C18:2- oder C18:3-Fettsäuren) entstehen so Fettsäuren, die sich von C18:2-Fettsäuren ableiten, wie GLA, DGLA oder ARA oder, die sich von C18:3-Fettsäuren ableiten, wie SDA, ETA oder EPA. Liegt in der für das Verfahren verwendeten Pflanze als ungesättigte Fettsäure nur Linolsäure (= LA, C18:2^{Δ^{9,12}}) vor, so können als Produkte des Verfahrens nur GLA, DGLA und ARA entstehen, die als freie Fettsäuren oder gebunden vorliegen können. Ist in der im Verfahren verwendeten Pflanze als ungesättigte Fettsäure nur α -Linolensäure (= ALA, C18:3^{Δ^{9,12,15}}) beispielsweise wie in Lein, so können als Produkte des Verfahrens nur SDA, ETA und EPA entstehen, die wie oben beschrieben als freie Fettsäuren oder gebunden vorliegen können. Durch Modifikation der Aktivität der an der Synthese beteiligten Enzyme Acyl-CoA:Lysophospholipid-Acyltransferase vorteilhaft in Kombination mit der Δ -5-, Δ -6-Desaturase und Δ -6-Elongase, oder der Δ -5-, Δ -8-Desaturase und Δ -9-Elongase oder in Kombination mit nur den ersten beiden Gene Δ -6-Desaturase und Δ -6-Elongase oder Δ -8-Desaturase und Δ -9-Elongase der Synthesekette lassen sich gezielt in den vorgenannten Organismen vorteilhaft in den vorgenannten Pflanzen nur einzelne Produkte herstellen. Durch die Aktivität der Δ -6-Desaturase und Δ -6-Elongase entstehen beispielsweise GLA und DGLA bzw. SDA und ETA, je nach Ausgangspflanze und ungesättigter Fettsäure. Bevorzugt entstehen DGLA bzw. ETA oder deren Mischungen. Wird die Δ -5-Desaturase zusätzlich in die Organismen vorteilhaft in die Pflanze eingebracht, so entstehen zusätzlich ARA oder EPA. Dies gilt auch für Organismen in die vorher die Δ -8-Desaturase und Δ -9-Elongase eingebracht wurde. Vorteilhaft werden nur ARA oder EPA oder deren Mischungen synthetisiert, abhängig von der in im Organismus bzw. in der Pflanze vorliegenden Fettsäure, die als Ausgangssubstanz für die Synthese dient. Da es sich um Biosyntheseketten handelt, liegen die jeweiligen Endprodukte nicht als Reinsubstanzen in den Organismen vor. Es sind immer auch geringe Mengen der Vorläuferverbindungen im Endprodukt enthalten. Diese geringen Mengen betragen weniger als 20 Gew.-%, vorteilhaft weniger als 15 Gew.-%, besonders vorteilhaft weniger als 10 Gew.-%, ganz besonders vorteilhaft weniger als 5, 4, 3, 2 oder 1 Gew.-% bezogen auf das Endprodukt DGLA, ETA oder deren Mischungen bzw. ARA, EPA oder deren Mischungen.

Zur Steigerung der Ausbeute im beschriebenen Verfahren zur Herstellung von Ölen und/oder Triglyceriden mit einem vorteilhaft erhöhten Gehalt an mehrfach ungesättigten Fettsäuren ist es vorteilhaft die Menge an Ausgangsprodukt für die Fettsäuresynthese zu steigern, dies kann beispielsweise durch das Einbringen einer Nukleinsäure in den Organismus, die für ein Polypeptid mit Δ -12-Desaturase codiert, erreicht werden. Dies ist besonders vorteilhaft in Öl-produzierenden Organismen wie Raps, die einen hohen Ölsäuregehalt aufweisen. Da diese Organismen nur einen geringen Gehalt an Linolsäure aufweisen (Mikoklajczak et al., Journal of the American Oil Chemical Society, 38, 1961, 678 - 681) ist die Verwendung der genannten Δ -12-Desaturasen zur Herstellung des Ausgangsprodukts Linolsäure vorteilhaft.

Im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren stammen vorteilhaft aus Pflanzen wie Algen wie *Isochrysis* oder *Cryptocodinium*, Algen/Diatomeen wie *Phaeodactylum*, Moose wie *Physcomitrella* oder *Ceratodon* oder höheren Pflanzen wie den *Primulaceae* wie *Aleuritia*, *Calendula stellata*, *Osteospermum spinescens* oder *Osteospermum hyoseroides*, Mikroorganismen wie Pilzen wie *Aspergillus*, *Thraustochytrium*, *Phytophthora*, *Entomophthora*, *Mucor* oder *Mortierella*, Hefen oder Tieren wie Nematoden wie *Caenorhabditis*, Insekten oder dem Mensch. Vorteilhaft stammen die Nukleinsäuren aus Pilzen, Tieren oder aus Pflanzen wie Algen oder Moosen, bevorzugt aus Nematoden wie *Caenorhabditis*.

10 Vorteilhaft werden im erfindungsgemäßen Verfahren die vorgenannten Nukleinsäuresequenzen oder deren Derivat oder Homologe, die für Polypeptide codieren, die noch die enzymatische Aktivität der durch Nukleinsäuresequenzen codierten Proteine besitzen. Diese Sequenzen werden einzeln oder in Kombination mit der für die Acyl-CoA:Lysophospholipid-Acyltransferase codierenden Nukleinsäuresequenz in Expressionskonstrukte cloniert und zum Einbringen und zur Expression in Organismen verwendet. Diese Expressionskonstrukte ermöglichen durch ihre Konstruktion eine vorteilhafte optimale Synthese der im erfindungsgemäßen Verfahren produzierten mehrfach ungesättigten Fettsäuren.

20 Bei einer bevorzugten Ausführungsform umfasst das Verfahren ferner den Schritt des Gewinnens einer Zelle oder eines ganzen Organismus, der die im Verfahren verwendeten Nukleinsäuresequenzen enthält, wobei die Zelle und/oder der Organismus mit der erfindungsgemäßen Nukleinsäuresequenz, die für die Acyl-CoA:Lysophospholipid-Acyltransferase codiert, einem Genkonstrukt oder einem Vektor wie nachfolgend beschrieben, allein oder in Kombination mit weiteren Nukleinsäuresequenzen, die für Proteine des Fettsäure- oder Lipidsstoffwechsels codieren, transformiert wird. Bei einer weiteren bevorzugten Ausführungsform umfasst dieses Verfahren ferner den Schritt des Gewinnens der Feinchemikalie aus der Kultur. Bei der Kultur kann es sich beispielsweise um eine Fermentationskultur beispielsweise im Falle der Kultivierung von Mikroorganismen wie z.B. *Mortierella*, *Saccharomyces* oder *Thraustochytrium* oder um eine Treibhaus oder Feldkultur einer Pflanze handeln. Die so hergestellte Zelle oder 30 der so hergestellte Organismus ist vorteilhaft eine Zelle eines Öl-produzierenden Organismus wie einer Ölfruchtpflanze wie beispielsweise Erdnuss, Raps, Canola, Lein, Hanf, Erdnuss, Soja, Safflower, Hanf, Sonnenblumen oder Borretsch.

35 Unter Anzucht ist beispielsweise die Kultivierung im Falle von Pflanzenzellen, -gewebe oder -organe auf oder in einem Nährmedium oder der ganzen Pflanze auf bzw. in einem Substrat beispielsweise in Hydrokultur, Blumentopferde oder auf einem Ackerboden zu verstehen.

40 "Transgen" bzw. "Rekombinant" im Sinne der Erfindung bedeutet bezüglich zum Beispiel einer Nukleinsäuresequenz, einer Expressionskassette (= Genkonstrukt) oder einem Vektor enthaltend die erfindungsgemäße Nukleinsäuresequenz oder einem Organismus transformiert mit den erfindungsgemäßen Nukleinsäuresequenzen, Expres-

sionskassette oder Vektor alle solche durch gentechnische Methoden zustandekommenen Konstruktionen, in denen sich entweder

- a) die erfindungsgemäße Nukleinsäuresequenz, oder
- b) eine mit der erfindungsgemäßen Nukleinsäuresequenz funktionell verknüpfte genetische Kontrollsequenz, zum Beispiel ein Promotor, oder
- c) (a) und (b)

sich nicht in ihrer natürlichen, genetischen Umgebung befinden oder durch gentechnische Methoden modifiziert wurden, wobei die Modifikation beispielhaft eine Substitution, Addition, Deletion, Inversion oder Insertion eines oder mehrerer Nukleotidreste sein kann. Natürliche genetische Umgebung meint den natürlichen genomischen bzw. chromosomalen Locus in dem Herkunftsorganismus oder das Vorliegen in einer genomischen Bibliothek. Im Fall einer genomischen Bibliothek ist die natürliche, genetische Umgebung der Nukleinsäuresequenz bevorzugt zumindest noch teilweise erhalten. Die Umgebung flankiert die Nukleinsäuresequenz zumindest an einer Seite und hat eine Sequenzlänge von mindestens 50 bp, bevorzugt mindestens 500 bp, besonders bevorzugt mindestens 1000 bp, ganz besonders bevorzugt mindestens 5000 bp. Eine natürlich vorkommende Expressionskassette - beispielsweise die natürlich vorkommende Kombination des natürlichen Promotors der erfindungsgemäßen Nukleinsäuresequenz mit dem entsprechenden Acyl-CoA:Lysophospholipid-Acyltransferase -Gen - wird zu einer transgenen Expressionskassette, wenn diese durch nicht-natürliche, synthetische ("künstliche") Verfahren wie beispielsweise einer Mutagenisierung geändert wird. Entsprechende Verfahren sind beispielsweise beschrieben in US 5,565,350 oder WO 00/15815.

Unter transgenen Organismus bzw. transgener Pflanze im Sinne der Erfindung ist wie vorgenannt zu verstehen, dass die im Verfahren verwendeten Nukleinsäuren nicht an ihrer natürlichen Stelle im Genom eines Organismus sind, dabei können die Nukleinsäuren homolog oder heterolog exprimiert werden. Transgen bedeutet aber auch wie genannt, dass die erfindungsgemäßen Nukleinsäuren an ihrem natürlichen Platz im Genom eines Organismus sind, dass jedoch die Sequenz gegenüber der natürlichen Sequenz verändert wurde und/oder dass die Regulationssequenzen, der natürlichen Sequenzen verändert wurden. Bevorzugt ist unter transgen die Expression der erfindungsgemäßen Nukleinsäuren an nicht natürlicher Stelle im Genom zu verstehen, das heißt eine homologe oder bevorzugt heterologe Expression der Nukleinsäuren liegt vor. Bevorzugte transgene Organismen sind Pilze wie *Mortierella* oder Pflanzen sind die Ölfruchtpflanzen.

Als Organismen bzw. Wirtsorganismen für die im erfindungsgemäßen Verfahren verwendeten Nukleinsäuren, die Expressionskassette oder den Vektor eignen sich prinzipiell vorteilhaft alle Organismen, die in der Lage sind Fettsäuren speziell ungesättigte Fettsäuren zu synthetisieren bzw. für die Expression rekombinanter Gene geeignet sind. Beispielhaft seien Pflanzen wie *Arabidopsis*, *Asteraceae* wie *Calendula* oder Kul-

5. turpflanzen wie Soja, Erdnuss, Rizinus, Sonnenblume, Mais, Baumwolle, Flachs, Raps, Kokosnuss, Ölpalme, Färbersafflor (*Carthamus tinctorius*) oder Kakaobohne, Mikroorganismen wie Pilze beispielsweise die Gattung *Mortierella*, *Saprolegnia* oder *Pythium*, Bakterien wie die Gattung *Escherichia*, Hefen wie die Gattung *Saccharomyces*, Cyanobakterien, Ciliaten, Algen oder Protozoen wie Dinoflagellaten wie *Cryptocodinium* genannt. Bevorzugt werden Organismen, die natürlicherweise Öle in größeren Mengen synthetisieren können wie Pilze wie *Mortierella alpina*, *Pythium insidiosum* oder Pflanzen wie Soja, Raps, Kokosnuss, Ölpalme, Färbersafflor, Flachs, Hanf, Rizinus, Calendula, Erdnuss, Kakaobohne oder Sonnenblume oder Hefen wie *Saccharomyces cerevisiae*, besonders bevorzugt werden Soja, Flachs, Raps, Färbersafflor, Sonnenblume, Calendula, *Mortierella* oder *Saccharomyces cerevisiae*. Prinzipiell sind als Wirtsorganismen neben den vorgenannten transgenen Organismen auch transgene Tiere vorteilhaft nicht-humane Tiere geeignet beispielsweise *C. elegans*.

15 Nutzbare Wirtszellen sind weiterhin genannt in: Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

Verwendbare Expressionsstämme z.B. solche, die eine geringere Proteaseaktivität aufweisen sind beschrieben in: Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128.

20 Hierzu gehören Pflanzenzellen und bestimmte Gewebe, Organe und Teile von Pflanzen in all ihren Erscheinungsformen, wie Antheren, Fasern, Wurzelhaare, Stängel, Embryos, Kalli, Kotelydonen, Petiolen, Erntematerial, pflanzliches Gewebe, reproduktives Gewebe und Zellkulturen, das von der eigentlichen transgenen Pflanze abgeleitet ist und/oder dazu verwendet werden kann, die transgene Pflanze hervorzubringen.

25 Transgene Pflanzen, die die im erfindungsgemäßen Verfahren synthetisierten mehrfach ungesättigten Fettsäuren enthalten, können vorteilhaft direkt vermarktet werden ohne dass, die synthetisierten Öle, Lipide oder Fettsäuren isoliert werden müssen. Unter Pflanzen im erfindungsgemäßen Verfahren sind ganze Pflanzen sowie alle Pflanzenteile, Pflanzenorgane oder Pflanzenteile wie Blatt, Stiel, Samen, Wurzel, Knollen, Antheren, Fasern, Wurzelhaare, Stängel, Embryos, Kalli, Kotelydonen, Petiolen, Erntematerial, pflanzliches Gewebe, reproduktives Gewebe, Zellkulturen, die sich von der transgenen Pflanze ableiten und/oder dazu verwendet werden können, die transgene Pflanze hervorzubringen. Der Samen umfasst dabei alle Samenteile wie die Samenhüllen, Epidermis- und Samenzellen, Endosperm oder Embryogewebe. Die im erfindungsgemäßen Verfahren hergestellten Verbindungen können aber auch aus den Organismen vorteilhaft Pflanzen in Form ihrer Öle, Fett, Lipide und/oder freien Fettsäuren isoliert werden. Durch dieses Verfahren hergestellte mehrfach ungesättigten Fettsäuren lassen sich durch Ernten der Organismen entweder aus der Kultur, in der sie wachsen, oder vom Feld ernten. Dies kann über Pressen oder Extraktion der Pflanzenteile bevorzugt der Pflanzensamen erfolgen. Dabei können die Öle, Fette, Lipide und/oder freien Fettsäuren durch sogenanntes kalt schlagen oder kalt pressen ohne Zuführung von Wärme durch Pressen gewonnen werden. Damit sich die Pflan-

30
35
40

zenteile speziell die Samen leichter aufschließen lassen, werden sie vorher zerkleinert, gedämpft oder geröstet. Die so vorbehandelten Samen können anschließend gepresst werden oder mit Lösungsmittel wie warmen Hexan extrahiert werden. Anschließend wird das Lösungsmittel wieder entfernt. Im Falle von Mikroorganismen werden diese nach Ernte beispielsweise direkt ohne weitere Arbeitsschritte extrahiert oder aber nach Aufschluss über verschiedene dem Fachmann bekannte Methoden extrahiert. Auf diese Weise können mehr als 96 % der im Verfahren hergestellten Verbindungen isoliert werden. Anschließend werden die so erhaltenen Produkte weiter bearbeitet, das heißt raffiniert. Dabei werden zunächst beispielsweise die Pflanzenschleime und Trübstoffe entfernt. Die sogenannte Entschleimung kann enzymatisch oder beispielsweise chemisch/physikalisch durch Zugabe von Säure wie Phosphorsäure erfolgen. Anschließend werden die freien Fettsäuren durch Behandlung mit einer Base beispielsweise Natronlauge entfernt. Das erhaltene Produkt wird zur Entfernung der im Produkt verbliebenen Lauge mit Wasser gründlich gewaschen und getrocknet. Um die noch im Produkt enthaltenen Farbstoffe zu entfernen werden die Produkte einer Bleichung mit beispielsweise Bleicherde oder Aktivkohle unterzogen. Zum Schluss wird das Produkt noch beispielsweise mit Wasserdampf noch desodoriert.

Vorzugsweise sind die durch dieses Verfahren produzierten PUFAs bzw. LCPUFAs C_{18} -, C_{20} - oder C_{22} -Fettsäuremoleküle mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise drei, vier, fünf oder sechs Doppelbindungen. Diese C_{18} -, C_{20} - oder C_{22} -Fettsäuremoleküle lassen sich aus dem Organismus in Form eines Öls, Lipids oder einer freien Fettsäure isolieren. Geeignete Organismen sind beispielsweise die vorstehend erwähnten. Bevorzugte Organismen sind transgene Pflanzen.

Eine Ausführungsform der Erfindung sind deshalb Öle, Lipide oder Fettsäuren oder Fraktionen davon, die durch das oben beschriebene Verfahren hergestellt worden sind, besonders bevorzugt Öl, Lipid oder eine Fettsäurezusammensetzung, die PUFAs umfassen und von transgenen Pflanzen herrühren.

Eine weitere erfindungsgemäße Ausführungsform ist die Verwendung des Öls, Lipids, der Fettsäuren und/oder der Fettsäurezusammensetzung in Futtermitteln, Nahrungsmitteln, Kosmetika oder Pharmazeutika.

Unter dem Begriff "Öl", "Lipid" oder "Fett" wird ein Fettsäuregemisch verstanden, das ungesättigte, gesättigte, vorzugsweise veresterte Fettsäure(n) enthält. Bevorzugt ist, dass das Öl, Lipid oder Fett einen hohen Anteil an mehrfach ungesättigten freien oder vorteilhaft veresterten Fettsäure(n), insbesondere Linolsäure, γ -Linolensäure, Dihomo- γ -linolensäure, Arachidonsäure, α -Linolensäure, Stearidonsäure, Eicosatetraensäure, Eicosapentaensäure, Docosapentaensäure oder Docosahexaensäure hat. Vorzugsweise ist der Anteil an ungesättigten veresterten Fettsäuren ungefähr 30 %, mehr bevorzugt ist ein Anteil von 50 %, noch mehr bevorzugt ist ein Anteil von 60 %, 70 %, 80 % oder mehr. Zur Bestimmung kann z.B. der Anteil an Fettsäure nach Überführung der Fettsäuren in die Methylester durch Umesterung gaschromatographisch bestimmt werden. Das Öl, Lipid oder Fett kann verschiedene andere gesättigte oder ungesättigte

Fettsäuren, z.B. Calendulasäure, Palmitin-, Palmitolein-, Stearin-, Ölsäure etc., enthalten. Insbesondere kann je nach Ausgangsorganismus der Anteil der verschiedenen Fettsäuren in dem Öl oder Fett schwanken.

5 Bei den im Verfahren hergestellten mehrfach ungesättigte Fettsäuren mit vorteilhaft mindestens zwei Doppelbindungen enthalten, handelt es sich beispielsweise um Sphingolipide, Phosphoglyceride, Lipide, Glycolipide, Phospholipide, Monoacylglycerin, Diacylglycerin, Triacylglycerin oder sonstige Fettsäureester.

10 Aus den so im erfindungsgemäßen Verfahren hergestellten mehrfach ungesättigte Fettsäuren mit vorteilhaft mindestens zwei Doppelbindungen lassen sich die enthalten- den mehrfach ungesättigten Fettsäuren beispielsweise über eine Alkalibehandlung beispielsweise wäßrige KOH oder NaOH oder saure Hydrolyse vorteilhaft in Gegen- wart eines Alkohols wie Methanol oder Ethanol oder über eine enzymatische Abspal- tung freisetzen und isolieren über beispielsweise Phasentrennung und anschließender Ansäuerung über z.B. H_2SO_4 . Die Freisetzung der Fettsäuren kann auch direkt ohne
15 die vorhergehend beschriebene Aufarbeitung erfolgen.

Die im Verfahren verwendeten Nukleinsäuren können nach Einbringung in einem Or- ganismus vorteilhaft einer Pflanzenzelle bzw. Pflanze entweder auf einem separaten Plasmid liegen oder in das Genom der Wirtszelle integriert sein. Bei Integration in das Genom kann die Integration zufallsgemäß sein oder durch derartige Rekombination
20 erfolgen, dass das native Gen durch die eingebrachte Kopie ersetzt wird, wodurch die Produktion der gewünschten Verbindung durch die Zelle moduliert wird, oder durch Verwendung eines Gens in trans, so dass das Gen mit einer funktionellen Expressi- onseinheit, welche mindestens eine die Expression eines Gens gewährleistende Se- quenz und mindestens eine die Polyadenylierung eines funktionell transkribierten Gens
25 gewährleistende Sequenz enthält, funktionell verbunden ist. Vorteilhaft werden die Nukleinsäuren über Multiexpressionskassetten oder Konstrukte zur multiparallelen Ex- pression in die Organismen vorteilhaft zur multiparallelen samenspezifischen Expressi- on von Genen in die Pflanzen gebracht.

30 Moose und Algen sind die einzigen bekannten Pflanzensysteme, die erhebliche Men- gen an mehrfach ungesättigten Fettsäuren, wie Arachidonsäure (ARA) und/oder Eico- sapentaensäure (EPA) und/oder Docosahexaensäure (DHA) herstellen. Moose enthal- ten PUFAs in Membranlipiden während Algen, algenverwandte Organismen und einige Pilze auch nennenswerte Mengen an PUFAs in der Triacylglycerolfraktion akkumulie- ren. Daher eignen sich Nukleinsäuremoleküle, die aus solchen Stämmen isoliert wer-
35 den, die PUFAs auch in der Triacylglycerolfraktion akkumulieren, besonders vorteilhaft für das erfindungsgemäße Verfahren und damit zur Modifikation des Lipid- und PUFA- Produktionssystems in einem Wirt, insbesondere Pflanzen, wie Ölfuchtpflanzen, bei- spielsweise Raps, Canola, Lein, Hanf, Soja, Sonnenblumen, Borretsch. Sie sind des- halb vorteilhaft im erfindungsgemäßen Verfahren verwendbar.

40 Als Substrate der im erfindungsgemäßen Acyl-CoA:Lysophospholipid-Acyltransferasen werden vorteilhaft C_{16} -, C_{18} -, C_{20} - oder C_{22} -Fettsäuren verwendet.

- Zur Herstellung der erfindungsgemäßen langkettiger PUFAs müssen die mehrfach ungesättigten C₁₆- oder C₁₈-Fettsäuren zunächst durch die enzymatische Aktivität einer Desaturase zunächst desaturiert und anschließend über eine Elongase um mindestens zwei Kohlenstoffatome verlängert werden. Nach einer Elongationsrunde führt diese Enzymaktivität zu C₁₈- oder C₂₀-Fettsäuren, und nach zwei oder drei Elongationsrunden zu C₂₂- oder C₂₄-Fettsäuren. Die Aktivität der erfindungsgemäßen Verfahren verwendeten Desaturasen und Elongasen führt vorzugsweise zu C₁₈-, C₂₀- und/oder C₂₂-Fettsäuren vorteilhaft mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit drei, vier oder fünf Doppelbindungen, besonders bevorzugt zu C₂₀- und/oder C₂₂-Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit drei, vier oder fünf Doppelbindungen im Molekül. Nachdem eine erste Desaturierung und die Verlängerung stattgefunden hat, können weitere Desaturierungsschritte wie z.B. eine solche in Δ-5-Position erfolgen. Besonders bevorzugt als Produkte des erfindungsgemäßen Verfahrens sind Dihomo-γ-linolensäure, Arachidonsäure, Eicosapentaensäure, Docosapentaensäure und/oder Docosahexaensäure. Die C₁₈-Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure können durch die erfindungsgemäße enzymatische Aktivität in Form der freien Fettsäure oder in Form der Ester, wie Phospholipide, Glycolipide, Sphingolipide, Phosphoglyceride, Monoacylglycerin, Diacylglycerin oder Triacylglycerin, verlängert werden.
- Der bevorzugte Biosyntheseort von Fettsäuren, Ölen, Lipiden oder Fette in den vorteilhaft verwendeten Pflanzen ist beispielsweise im allgemeinen der Samen oder Zellschichten des Samens, so dass eine sämenspezifische Expression der im Verfahren verwendeten Nukleinsäuren sinnvoll ist. Es ist jedoch naheliegend, dass die Biosynthese von Fettsäuren, Ölen oder Lipiden nicht auf das Samengewebe beschränkt sein muss, sondern auch in allen übrigen Teilen der Pflanze - beispielsweise in Epidermiszellen oder in den Knollen - gewebespezifisch erfolgen kann.
- Werden im erfindungsgemäßen Verfahren als Organismen Mikroorganismus wie Hefen wie *Saccharomyces* oder *Schizosaccharomyces*, Pilze wie *Mortierella*, *Aspergillus*, *Phytophthora*, *Entomophthora*, *Mucor* oder *Traustochytrium* Algen wie *Isochrysis*, *Phaeodactylum* oder *Cryptocodinium* verwendet, so werden diese Organismen vorteilhaft fermentativ angezogen.
- Durch die Verwendung der erfindungsgemäßen Nukleinsäuren, die für Acyl-CoA:Lysophospholipid-Acyltransferasen codieren, können im Verfahren die hergestellten mehrfach ungesättigten Fettsäuren mindestens um 10 %, bevorzugt mindestens um 15 %, besonders bevorzugt mindestens um 20 %, ganz besonders bevorzugt um mindestens 50 % gegenüber dem Wildtyp der Organismen, die die Nukleinsäuren nicht rekombinant enthalten, erhöht werden.
- Durch das erfindungsgemäße Verfahren können die hergestellten mehrfach ungesättigten Fettsäuren in den im Verfahren verwendeten Organismen prinzipiell auf zwei Arten erhöht werden. Es kann vorteilhaft der Pool an freien mehrfach ungesättigten Fettsäuren und/oder der Anteil der über das Verfahren hergestellten veresterten mehr-

fach ungesättigten Fettsäuren erhöht werden. Vorteilhaft wird durch das erfindungsgemäße Verfahren der Pool an veresterten mehrfach ungesättigten Fettsäuren in den transgenen Organismen erhöht.

- 5 Werden im erfindungsgemäßen Verfahren als Organismen Mikroorganismen verwendet, so werden sie je nach Wirtsorganismus in dem Fachmann bekannter Weise angezogen bzw. gezüchtet. Mikroorganismen werden in der Regel in einem flüssigen Medium, das eine Kohlenstoffquelle meist in Form von Zuckern, eine Stickstoffquelle meist in Form von organischen Stickstoffquellen wie Hefeextrakt oder Salzen wie Ammoniumsulfat, Spurenelemente wie Eisen-, Mangan-, Magnesiumsalze und gegebenenfalls
10 Vitamine enthält, bei Temperaturen zwischen 0 °C und 100 °C, bevorzugt zwischen 10 °C bis 60 °C unter Sauerstoffbegasung angezogen. Dabei kann der pH der Nährflüssigkeit auf einen festen Wert gehalten werden, das heißt während der Anzucht reguliert werden oder nicht. Die Anzucht kann batch weise, semi batch weise oder kontinuierlich erfolgen. Nährstoffe können zu Beginn der Fermentation vorgelegt oder semi-
15 kontinuierlich oder kontinuierlich nachgefüttert werden. Die hergestellten mehrfach ungesättigten Fettsäuren können nach dem Fachmann bekannten Verfahren wie oben beschrieben aus den Organismen isoliert werden. Beispielsweise über Extraktion, Destillation, Kristallisation, ggf. Salzfällung und/oder Chromatographie. Die Organismen können dazu vorher noch vorteilhaft aufgeschlossen werden.

- 20 Das erfindungsgemäße Verfahren wird, wenn es sich bei den Wirtsorganismen um Mikroorganismen handelt, vorteilhaft bei einer Temperatur zwischen 0 °C bis 95 °C, bevorzugt zwischen 10 °C bis 85 °C, besonders bevorzugt zwischen 15 °C bis 75 °C, ganz besonders bevorzugt zwischen 15 °C bis 45 °C durchgeführt

- 25 Der pH-Wert wird dabei vorteilhaft zwischen pH 4 und 12, bevorzugt zwischen pH 6 und 9, besonders bevorzugt zwischen pH 7 und 8 gehalten.

- Das erfindungsgemäße Verfahren kann batchweise, semi-batchweise oder kontinuierlich betrieben werden. Eine Zusammenfassung über bekannte Kultivierungsmethoden ist im Lehrbuch von Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) oder im Lehrbuch von Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)) zu
30 finden.

- Das zu verwendende Kulturmedium hat in geeigneter Weise den Ansprüchen der jeweiligen Stämme zu genügen. Beschreibungen von Kulturmedien verschiedener Mikroorganismen sind im Handbuch "Manual of Methods für General Bacteriology" der
35 American Society für Bacteriology (Washington D. C., USA, 1981) enthalten.

- 7 Diese erfindungsgemäß einsetzbaren Medien umfassen wie oben beschrieben gewöhnlich eine oder mehrere Kohlenstoffquellen, Stickstoffquellen, anorganische Salze, Vitamine und/oder Spurenelemente.

Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind beispielsweise Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cel-
5, lulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte der Zucker-Raffinierung zu den Medien geben. Es kann auch vor-
10 teilhaft sein, Gemische verschiedener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Öle und Fette wie z. B. Sojaöl, Sonnenblumenöl, Erdnussöl und/oder Kokosfett, Fettsäuren wie z. B. Palmitinsäure, Stearinsäure und/oder Linol-
säure, Alkohole und/oder Polyalkohole wie z. B. Glycerin, Methanol und/oder Ethanol und/oder organische Säuren wie z. B. Essigsäure und/oder Milchsäure.

Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffver-
bindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stick-
stoffquellen umfassen Ammoniak in flüssiger- oder gasform oder Ammoniumsalze, wie
15 Ammoniumsulfat, Ammoniumchlorid, Ammoniumphosphat, Ammoniumcarbonat oder Ammoniumnitrat, Nitrate, Harnstoff, Aminosäuren oder komplexe Stickstoffquellen, wie Maisquellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und andere. Die Stickstoffquellen können einzeln oder als Mischung verwendet werden.

Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen
20 die Chlorid-, Phosphor- oder Sulfatsalze von Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen.

Als Schwefelquelle für die Herstellung von schwefelhaltigen Feinchemikalien, insbe-
sondere von Methionin, können anorganische schwefelhaltige Verbindungen wie bei-
spielsweise Sulfate, Sulfite, Dithionite, Tetrathionate, Thiosulfate, Sulfide aber auch
organische Schwefelverbindungen, wie Mercaptane und Thiole, verwendet werden.

25 Als Phosphorquelle können Phosphorsäure, Kaliumdihydrogenphosphat oder Dikali-
umhydrogenphosphat oder die entsprechenden Natrium haltigen Salze verwendet
werden.

Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu
halten. Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol
30 oder Protocatechuat, oder organische Säuren, wie Citronensäure.

Die erfindungsgemäß zur Kultivierung von Mikroorganismen eingesetzten Fermentati-
onsmedien enthalten üblicherweise auch andere Wachstumsfaktoren, wie Vitamine
oder Wachstumsförderer, zu denen beispielsweise Biotin, Riboflavin, Thiamin, Folsäu-
re, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze
35 stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Melassen,
Maisquellwasser und dergleichen. Dem Kulturmedium können überdies geeignete Vor-
stufen zugesetzt werden. Die genaue Zusammensetzung der Medienverbindungen
hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individu-
ell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehr-
40 buch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F.

Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wachstumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.

5 Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der Anzucht zugegeben sein oder wahlfrei kontinuierlich oder chargenweise hinzugegeben werden.

10 Die Temperatur der Kultur liegt normalerweise zwischen 15°C und 45°C, vorzugsweise bei 25°C bis 40°C und kann während des Experimentes konstant gehalten oder verändert werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen. Der pH-Wert für die Anzucht lässt sich während der Anzucht durch Zugabe von basischen Verbindungen wie Natriumhydroxid, Kaliumhydroxid, Ammoniak bzw. Ammoniakwasser oder sauren Verbindungen wie Phosphorsäure oder Schwefelsäure kontrollieren. Zur Kontrolle der Schaumentwicklung können Antischaummittel wie z. B. Fettsäurepolyglykolester, eingesetzt werden. Zur Aufrechterhaltung der Stabilität von Plasmiden können dem Medium geeignete selektiv wirkende Stoffe, wie z. B. Antibiotika, hinzugefügt werden. Um aerobe Bedingungen aufrechtzuerhalten, werden Sauerstoff oder Sauerstoff haltige Gasgemischungen, wie z. B. Umgebungsluft, in die Kultur eingetragen. Die Temperatur der Kultur liegt normalerweise bei 20°C bis 45°C und vorzugsweise bei 25°C bis 40°C. Die Kultur wird solange fortgesetzt, bis sich ein Maximum des gewünschten Produktes gebildet hat. Dieses Ziel wird normalerweise innerhalb von 10 Stunden bis 160 Stunden erreicht.

25 Die so erhaltenen, insbesondere mehrfach ungesättigte Fettsäuren enthaltenden, Fermentationsbrühen haben üblicherweise eine Trockenmasse von 7,5 bis 25 Gew.-%.

30 Die Fermentationsbrühe kann anschließend weiterverarbeitet werden. Je nach Anforderung kann die Biomasse ganz oder teilweise durch Separationsmethoden, wie z. B. Zentrifugation, Filtration, Dekantieren oder einer Kombination dieser Methoden aus der Fermentationsbrühe entfernt oder vollständig in ihr belassen werden. Vorteilhaft wird die Biomasse nach Abtrennung aufgearbeitet.

35 Die Fermentationsbrühe kann aber auch ohne Zellabtrennung mit bekannten Methoden, wie z. B. mit Hilfe eines Rotationsverdampfers, Dünnschichtverdampfers, Fallfilmverdampfers, durch Umkehrosmose, oder durch Nanofiltration, eingedickt beziehungsweise aufkonzentriert werden. Diese aufkonzentrierte Fermentationsbrühe kann schließlich zur Gewinnung der darin enthaltenen Fettsäuren aufgearbeitet werden.

Die im Verfahren gewonnenen Fettsäuren eignen sich auch als Ausgangsmaterial für die chemische Synthese von weiteren Wertprodukten. Sie können beispielsweise in Kombination miteinander oder allein zur Herstellung von Pharmaka, Nahrungsmittel, Tierfutter oder Kosmetika verwendet werden.

Ein weiterer erfindungsgemäßer Gegenstand sind isolierte Nukleinsäuresequenzen, die für Polypeptide mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codieren, wobei die durch die Nukleinsäuresequenzen codierten Acyl-CoA:Lysophospholipid-Acyltransferasen spezifisch C₁₆-, C₁₈-, C₂₀- oder C₂₂-Fettsäuren mit mindestens einer Doppelbindung im Fettsäuremolekül umsetzen.

Vorteilhafte isolierte Nukleinsäuresequenzen sind Sequenzen ausgewählt aus der Gruppe:

- a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Sequenz,
- 10 b) Nukleinsäuresequenzen, die sich als Ergebnis des degenerierten genetischen Codes von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 enthaltenden codierenden Sequenz ableiten lassen
- c) 15 Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 dargestellten Aminosäuresequenz codieren und mindestens 40 % Homologie auf Aminosäureebene mit SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 aufweisen und eine Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität aufweisen.
- 20 Vorteilhaft stammen die oben genannten Nukleinsäuresequenzen aus einem eukaryontischen Organismus.

Die im Verfahren verwendeten Nukleinsäuresequenzen, die für Proteine mit Acyl-CoA:Lysophospholipid-Acyltransferase-Aktivität codieren oder für Proteine des Fettsäure- oder Lipidstoffwechsels, werden vorteilhaft in einer Expressionskassette (= Nukleinsäurekonstrukt), die die Expression der Nukleinsäuren in einem Organismus vorteilhaft einer Pflanze oder einem Mikroorganismus ermöglicht, eingebracht.

30 Zum Einbringen werden die im Verfahren verwendeten Nukleinsäuren vorteilhaft einer Amplifikation und Ligation in bekannter Weise unterworfen. Vorzugsweise geht man in Anlehnung an das Protokoll der Pfu-DNA-Polymerase oder eines Pfu/Taq-DNA-Polymerasegemisches vor. Die Primer werden in Anlehnung an die zu amplifizierende Sequenz gewählt. Zweckmäßigerweise sollten die Primer so gewählt werden, dass das Amplifikat die gesamte kodogene Sequenz vom Start- bis zum Stop-Kodon umfasst. Im Anschluss an die Amplifikation wird das Amplifikat zweckmäßigerweise analysiert. Beispielsweise kann die Analyse nach gelelektrophoretischer Auftrennung hinsichtlich 35 Qualität und Quantität erfolgen. Im Anschluss kann das Amplifikat nach einem Standardprotokoll gereinigt werden (z.B. Qiagen). Ein Aliquot des gereinigten Amplifikats steht dann für die nachfolgende Klonierung zur Verfügung. Geeignete Klonierungsvektoren sind dem Fachmann allgemein bekannt. Hierzu gehören insbesondere Vektoren, die in mikrobiellen Systemen replizierbar sind, also vor allem Vektoren, die eine effi-

ziente Klonierung in Hefen oder Pilze gewährleisten, und die stabile Transformation von Pflanzen ermöglichen. Zu nennen sind insbesondere verschiedene für die T-DNA-vermittelte Transformation geeignete, binäre und co-integrierte Vektorsysteme. Derartige Vektorsysteme sind in der Regel dadurch gekennzeichnet, dass sie zumindest die für die Agrobakterium-vermittelte Transformation benötigten *vir*-Gene sowie die T-DNA begrenzenden Sequenzen (T-DNA-Border) beinhalten. Vorzugsweise umfassen diese Vektorsysteme auch weitere cis-regulatorische Regionen wie Promotoren und Terminatoren und/oder Selektionsmarker, mit denen entsprechend transformierte Organismen identifiziert werden können. Während bei co-integrierten Vektorsystemen *vir*-Gene und T-DNA-Sequenzen auf demselben Vektor angeordnet sind, basieren binäre Systeme auf wenigstens zwei Vektoren, von denen einer *vir*-Gene, aber keine T-DNA und ein zweiter T-DNA, jedoch kein *vir*-Gen trägt. Dadurch sind letztere Vektoren relativ klein, leicht zu manipulieren und sowohl in *E. coli* als auch in *Agrobacterium* zu replizieren. Zu diesen binären Vektoren gehören Vektoren der Serien pBIB-HYG, pPZP, pBecks, pGreen. Erfindungsgemäß bevorzugt verwendet werden Bin19, pBI101, pBinAR, pGPTV und pCAMBIA. Eine Übersicht über binäre Vektoren und ihre Verwendung gibt Hellens et al, Trends in Plant Science (2000) 5, 446-451. Für die Vektorpräparation können die Vektoren zunächst mit Restriktionsendonuklease(n) linearisiert und dann in geeigneter Weise enzymatisch modifiziert werden. Im Anschluss wird der Vektor gereinigt und ein Aliquot für die Klonierung eingesetzt. Bei der Klonierung wird das enzymatisch geschnittenen und erforderlichenfalls gereinigten Amplifikat mit ähnlich präparierten Vektorfragmenten mit Einsatz von Ligase kloniert. Dabei kann ein bestimmtes Nukleinsäurekonstrukt bzw. Vektor- oder Plasmidkonstrukt einen oder auch mehrere kodogene Genabschnitte aufweisen. Vorzugsweise sind die kodogenen Genabschnitte in diesen Konstrukten mit regulatorischen Sequenzen funktional verknüpft. Zu den regulatorischen Sequenzen gehören insbesondere pflanzliche Sequenzen wie die oben beschriebenen Promotoren und Terminatoren. Die Konstrukte lassen sich vorteilhafterweise in Mikroorganismen, insbesondere *Escherichia coli* und *Agrobacterium tumefaciens*, unter selektiven Bedingungen stabil propagieren und ermöglichen einen Transfer von heterologer DNA in Pflanzen oder Mikroorganismen.

Unter der vorteilhaften Verwendung von Klonierungsvektoren können die im Verfahren verwendeten Nukleinsäuren, die erfinderischen Nukleinsäuren und Nukleinsäurekonstrukte in Organismen wie Mikroorganismen oder vorteilhaft Pflanzen eingebracht werden und damit bei der Pflanzentransformation verwendet werden, wie denjenigen, die veröffentlicht sind in und dort zitiert sind: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Kapitel 6/7, S. 71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, 15-38; B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225)). Die im Verfahren verwendeten Nukleinsäuren, die erfinderischen Nukleinsäuren und Nukleinsäurekonstrukte und/oder Vektoren lassen sich damit zur gentechnologischen Veränderung

eines breiten Spektrums an Organismen vorteilhaft an Pflanzen verwenden, so dass diese bessere und/oder effizientere Produzenten von PUFAs werden.

Es gibt eine Reihe von Mechanismen, durch die die Veränderung eines erfindungsgemäßen Acyl-CoA:Lysophospholipid-Acyltransferase-Proteins die Ausbeute, Produktion und/oder Effizienz der Produktion einer Feinchemikalie aus einer Ölfuchtpflanze oder einem Mikroorganismus aufgrund eines veränderten Proteins direkt beeinflussen kann. Die Anzahl oder Aktivität des Acyl-CoA:Lysophospholipid-Acyltransferase-Proteins oder -Gens sowie von Genkombinationen von Acyl-CoA:Lysophospholipid-Acyltransferasen, Desaturasen und/oder Elongasen kann erhöht sein, so dass größere Mengen der produzierten Verbindungen de novo hergestellt werden, weil den Organismen diese Aktivität und Fähigkeit zur Biosynthese vor dem Einbringen des/der entsprechenden Gens/Gene fehlte. Entsprechendes gilt für die Kombination mit weiteren Desaturasen oder Elongasen oder weiteren Enzymen aus dem Fettsäure- und Lipidstoffwechsel. Auch die Verwendung verschiedener divergenter, d.h. auf DNA-Sequenzebene unterschiedlicher Sequenzen kann dabei vorteilhaft sein bzw. die Verwendung von Promotoren zur Genexpression, die eine andere zeitliche Genexpression z.B. abhängig vom Reifegrad eines Samens oder Öl-speichernden Gewebes ermöglicht.

Durch das Einbringen eines Acyl-CoA:Lysophospholipid-Acyltransferase-, Desaturase- und/oder Elongase-Gens oder mehrerer Acyl-CoA:Lysophospholipid-Acyltransferasen-, Desaturase- und/oder Elongase-Gene in einen Organismus allein oder in Kombination mit anderen Genen in eine Zelle kann nicht nur den Biosynthesefluss zum Endprodukt erhöht, sondern auch die entsprechende Triacylglycerin-Zusammensetzung erhöht oder de novo geschaffen werden. Ebenso kann die Anzahl oder Aktivität anderer Gene, die am Import von Nährstoffen, die zur Biosynthese einer oder mehrerer Feinchemikalien (z.B. Fettsäuren, polaren und neutralen Lipiden) nötig sind, erhöht sein, so dass die Konzentration dieser Vorläufer, Cofaktoren oder Zwischenverbindungen innerhalb der Zellen oder innerhalb des Speicherkompartiments erhöht ist, wodurch die Fähigkeit der Zellen zur Produktion von PUFAs, wie im folgenden beschrieben, weiter gesteigert wird. Fettsäuren und Lipide sind selbst als Feinchemikalien wünschenswert; durch Optimierung der Aktivität oder Erhöhung der Anzahl einer oder mehrerer Acyl-CoA:Lysophospholipid-Acyltransferase-, Desaturasen und/oder Elongasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Zerstören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann es möglich sein, die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmolekülen aus Organismen und vorteilhaft aus Pflanzen zu steigern.

Die im erfindungsgemäßen Verfahren verwendeten isolierten Nukleinsäuremoleküle codieren für Proteine oder Teile von diesen, wobei die Proteine oder das einzelne Protein oder Teile davon eine Aminosäuresequenz enthält, die ausreichend homolog zu einer Aminosäuresequenz der Sequenz SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 ist, so dass das Protein oder der Teil davon eine Acyl-

CoA:Lysophospholipid-Acyltransferase-Aktivität beibehält. Vorzugsweise hat das Protein oder der Teil davon, das/der von dem Nukleinsäuremolekül kodiert wird, noch seine wesentliche enzymatische Aktivität und die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen oder Lipidkörperchen in Organismen vorteilhaft in Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilzunehmen noch hat. Vorteilhaft ist das von den Nukleinsäuremolekülen kodierte Protein zu mindestens etwa 40 %, vorzugsweise mindestens etwa 60 % und stärker bevorzugt mindestens etwa 70 %, 80 % oder 90 % und am stärksten bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr homolog zu einer Aminosäuresequenz der Sequenz SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8. Vorteilhafte Ausführungsformen der erfindungsgemäßen Aminosäuresequenz der Sequenz SEQ ID NO: 2 sind Aminosäuresequenzen, die an Position 30 der SEQ ID NO: 2 anstelle des vorhandenen Methionin einen Valinrest haben oder in Position 100 anstelle des vorhandenen Serin einen Glycinrest haben oder in Position 170 anstelle des vorhandenen Phenylalanin einen Serinrest haben. Diese werden in SEQ ID NO: 4, SEQ ID NO: 6 bzw. SEQ ID NO: 8 wiedergegeben. Im Sinne der Erfindung ist unter Homologie oder homolog, Identität oder identisch zu verstehen.

Unter wesentlicher enzymatischer Aktivität der verwendeten Acyl-CoA:Lysophospholipid-Acyltransferasen ist zu verstehen, dass sie gegenüber den durch die Sequenz mit SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 und deren Derivate codierten Proteinen/Enzymen im Vergleich noch mindestens eine enzymatische Aktivität von mindestens 10 %, bevorzugt 20 %, besonders bevorzugt 30 % und ganz besonders 40 % aufweisen und damit am Stoffwechsel von zum Aufbau von Fettsäuren in einem Organismus vorteilhaft einer Pflanzenzelle notwendigen Verbindungen oder am Transport von Molekülen über Membranen teilnehmen können, wobei desaturierte C16-, C18- oder C20-24-Kohlenstoffketten mit Doppelbindungen an mindestens zwei, vorteilhaft drei, vier oder fünf Stellen gemeint ist.

Vorteilhaft im Verfahren verwendbare Nukleinsäuren stammen aus Pilzen oder Pflanzen wie Algen oder Moosen wie den Gattungen *Physcomitrella*, *Thraustochytrium*, *Phytophthora*, *Ceratodon*, *Isochrysis*, *Aleurita*, *Muscarioides*, *Mortierella*, *Borago*, *Phaeodactylum*, *Cryptothecodinium* oder aus Nematoden wie *Caenorhabditis*, speziell aus den Gattungen und Arten *Physcomitrella patens*, *Phytophthora infestans*, *Ceratodon purpureus*, *Isochrysis galbana*, *Aleurita farinosa*, *Muscarioides viallii*, *Mortierella alpina*, *Borago officinalis*, *Phaeodactylum tricornutum* oder besonders vorteilhaft aus *Caenorhabditis elegans*.

Alternativ können die verwendeten isolierten Nukleotidsequenzen für Acyl-CoA:Lysophospholipid-Acyltransferasen codieren, die an eine Nukleotidsequenz der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 hybridisieren, z.B. unter stringenten Bedingungen hybridisieren.

Die im Verfahren verwendeten Nukleinsäuresequenzen werden vorteilhaft in einer Expressionskassette, die die Expression der Nukleinsäuren in Organismen wie Mikroorganismen oder Pflanzen ermöglicht, eingebracht.

5 Dabei werden die für die erfinderische Acyl-CoA:Lysophospholipid-Acyltransferasen, die verwendeten Desaturasen und/oder die Elongasen codierenden Nukleinsäuresequenzen mit einem oder mehreren Regulationssignalen vorteilhafterweise zur Erhöhung der Genexpression funktionell verknüpft. Diese regulatorischen Sequenzen sollen die gezielte Expression der Gene und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion 10 exprimiert und/oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird. Beispielsweise handelt es sich bei diesen regulatorischen Sequenzen um Sequenzen an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert 15 worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Die Expressionskassette (= Expressionskonstrukt = Genkonstrukt) kann aber auch einfacher aufgebaut sein, das heißt es wurden keine zusätzlichen Regulationssignale vor die Nukleinsäuresequenz oder dessen Derivate inseriert 20 und der natürliche Promotor mit seiner Regulation wurde nicht entfernt. Stattdessen wurde die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und/oder die Genexpression gesteigert wird. Diese veränderten Promotoren können in Form von Teilsequenzen (= Promotor mit Teilen der erfindungsgemäßen Nukleinsäuresequenzen) auch allein vor das natürliche Gen zur Steigerung der Aktivität 25 gebracht werden. Das Genkonstrukt kann außerdem vorteilhafterweise auch eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die Acyl- 30 CoA:Lysophospholipid-Acyltransferase-Gene sowie die vorteilhaft verwendeten Δ -4-Desaturase-, Δ 5-Desaturase-, Δ -6-Desaturase- und/oder Δ -8-Desaturase-Gene und/oder die Δ -5-Elongase-, Δ -6-Elongase- und/oder Δ -9-Elongase-Gene können in einer oder mehreren Kopien in der Expressionskassette (= Genkonstrukt) enthalten sein. Vorteilhaft liegt nur jeweils eine Kopie der Gene in der Expressionskassette vor. 35 Dieses Genkonstrukt oder die Genkonstrukte können zusammen im Wirtsorganismus exprimiert werden. Dabei kann das Genkonstrukt oder die Genkonstrukte in einem oder mehreren Vektoren inseriert sein und frei in der Zelle vorliegen oder aber im Genom inseriert sein. Es ist vorteilhaft für die Insertion weiterer Gene im Wirtsgenom, wenn die zu exprimierenden Gene zusammen in einem Genkonstrukt vorliegen.

40 Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie

Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

- 5 Eine weitere Ausführungsform der Erfindung sind ein oder mehrere Genkonstrukte, die eine oder mehrere Sequenzen enthalten, die durch SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 oder dessen Derivate definiert sind und für Polypeptide gemäß SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 kodieren. Die genannten Acyl-CoA:Lysophospholipid-Acyltransferasen führen dabei zu einem Austausch der Fettsäuren zwischen dem Mono-, Di- und/oder Triglyceridpool der Zelle und dem CoA-Fettsäureester-Pool, wobei das Substrat vorteilhaft ein, zwei, drei, vier oder 10 fünf Doppelbindungen aufweist und vorteilhaft 16, 18, 20, 22 oder 24 Kohlenstoffatome im Fettsäuremolekül aufweist. Gleiches gilt für ihre Homologen, Derivate oder Analoga, die funktionsfähig mit einem oder mehreren Regulationssignalen, vorteilhafterweise zur Steigerung der Genexpression, verbunden sind.
- 15 Vorteilhafte Regulationssequenzen für das neue Verfahren liegen beispielsweise in Promotoren vor, wie dem *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacIq*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, λ -PR- oder λ -PL-Promotor und werden vorteilhafterweise in Gram-negativen Bakterien angewendet. Weitere vorteilhafte Regulationssequenzen liegen beispielsweise in den Gram-positiven Promotoren *amy* und *SPO2*, in den Hefe- oder Pilzpromotoren *ADC1*, *MF α* , *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH* oder in den Pflanzenpromotoren *CaMV/35S* [Franck et al., Cell 21 (1980) 285–294], *PRP1* [Ward et al., Plant. Mol. Biol. 22 (1993)], *SSU*, *OCS*, *lib4*, *usp*, *STLS1*, *B33*, *nos* oder im Ubiquitin- oder Phaseolin-Promotor vor. In diesem Zusammenhang vorteilhaft sind ebenfalls induzierbare Promotoren, wie die in EP-A-0 388 186 (Benzylsulfonamid-induzierbar), Plant J. 2, 1992:397–404 (Gatz et al., Tetracyclin-induzierbar), EP-A-0 335 528 (Abzisinäure-induzierbar) oder WO 93/21334 (Ethanol- oder Cyclohexenol-induzierbar) beschriebenen Promotoren. Weitere geeignete Pflanzenpromotoren sind der Promotor von cytosolischer FBPase oder der ST-LSI-Promotor der Kartoffel (Stockhaus et al., EMBO J. 8, 1989, 2445), der Phosphoribosylpyrophosphatamidotransferase-Promotor aus Glycine max (Genbank-Zugangsnr. U87999) oder der in EP-A-0 249 676 beschriebene nodienspezifische Promotor. Besonders vorteilhafte Promotoren sind Promotoren, welche die Expression in Geweben ermöglichen, die an der Fettsäurebiosynthese beteiligt sind. Ganz besonders vorteilhaft sind samenspezifische Promotoren, wie der ausführungsgemäße USP Promotor aber 30 auch andere Promotoren wie der *LeB4*-, *DC3*, Phaseolin- oder Napin-Promotor. Weitere besonders vorteilhafte Promotoren sind samenspezifische Promotoren, die für monokotyle oder dikotyle Pflanzen verwendet werden können und in US 5,608,152 (Napin-Promotor aus Raps), WO 98/45461 (Oleosin-Promotor aus Arabidopsis), US 5,504,200 (Phaseolin-Promotor aus Phaseolus vulgaris), WO 91/13980 (Bce4-Promotor aus Brassica), von Bäumlein et al., Plant J., 2, 2, 1992:233–239 (*LeB4*-Promotor aus einer Leguminose) beschrieben sind, wobei sich diese Promotoren für Dikotyledonen eignen. Die folgenden Promotoren eignen sich beispielsweise für Monokotyledonen *lpt-2*- oder *lpt-1*-Promotor aus Gerste (WO 95/15389 und 40

WO 95/23230), Hordein-Promotor aus Gerste und andere, in WO 99/16890 beschriebene geeignete Promotoren.

Es ist im Prinzip möglich, alle natürlichen Promotoren mit ihren Regulationssequenzen, wie die oben genannten, für das neue Verfahren zu verwenden. Es ist ebenfalls möglich und vorteilhaft, zusätzlich oder alleine synthetische Promotoren zu verwenden, besonders wenn sie eine Samen-spezifische Expression vermitteln, wie z.B. beschrieben in WO 99/16890.

Um einen besonders hohen Gehalt an PUFAs vor allem in transgenen Pflanzen zu erzielen, sollten die PUFA-Biosynthesegene vorteilhaft samenspezifisch in Ölsaaten exprimiert werden. Hierzu können Samen-spezifische Promotoren verwendet werden, bzw. solche Promotoren die im Embryo und/oder im Endosperm aktiv sind. Samen-spezifische Promotoren können prinzipiell sowohl aus dikotyledonen als auch aus monokotyledonen Pflanzen isoliert werden. Im folgenden sind vorteilhafte bevorzugte Promotoren aufgeführt: USP (= unknown seed protein) und Vicilin (*Vicia faba*) [Bäumlein et al., Mol. Gen. Genet., 1991, 225(3)], Napin (Raps) [US 5,608,152], Acyl-Carrier Protein (Raps) [US 5,315,001 und WO 92/18634], Oleosin (*Arabidopsis thaliana*) [WO 98/45461 und WO 93/20216], Phaseolin (*Phaseolus vulgaris*) [US 5,504,200], Bce4 [WO 91/13980], Leguminosen B4 (LegB4-Promotor) [Bäumlein et al., Plant J., 2,2, 1992], Lpt2 und Lpt1 (Gerste) [WO 95/15389 u. WO 95/23230], Samen-spezifische Promotoren aus Reis, Mais u. Weizen [WO 99/16890], Amy32b, Amy 6-6 und Aleurain [US 5,677,474], Bce4 (Raps) [US 5,530,149], Glycinin (Soja) [EP 571 741], Phosphoenol-Pyruvatcarboxylase (Soja) [JP 06/62870], ADR12-2 (Soja) [WO 98/08962], Isocitratlyase (Raps) [US 5,689,040] oder α -Amylase (Gerste) [EP 781 849].

Die Pflanzengenexpression lässt sich auch über einen chemisch induzierbaren Promotor erleichtern (siehe eine Übersicht in Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108). Chemisch induzierbare Promotoren eignen sich besonders, wenn gewünscht wird, dass die Genexpression auf zeitspezifische Weise erfolgt. Beispiele für solche Promotoren sind ein Salicylsäure-induzierbarer Promotor (WO 95/19443), ein Tetracyclin-induzierbarer Promotor (Gatz et al. (1992) Plant J. 2, 397-404) und ein Ethanol-induzierbarer Promotor.

Um eine stabile Integration der Biosynthesegene in die transgene Pflanze über mehrere Generation sicherzustellen, sollte jede der im Verfahren verwendeten Nukleinsäuren, die für die Acyl-CoA:Lysophospholipid-Acyltransferase, die vorteilhafte Δ -4-Desaturase, die Δ -5-Desaturase, die Δ -6-Desaturase, die Δ -8-Desaturase und/oder die Δ -5-Elongase, die Δ -6-Elongase und/oder die Δ -9-Elongase codieren, unter der Kontrolle eines eigenen bevorzugt eines unterschiedlichen Promotors exprimiert werden, da sich wiederholende Sequenzmotive zu Instabilität der T-DNA bzw. zu Rekombinationsereignissen führen können. Die Expressionskassette ist dabei vorteilhaft so aufgebaut, dass einem Promotor eine geeignete Schnittstelle zur Insertion der zu exprimierenden Nukleinsäure folgt vorteilhaft in einem Polylinker anschließend gegebenenfalls ein Terminator hinter dem Polylinker liegt. Diese Abfolge wiederholt sich mehrfach be-

vorzugt drei-, vier- oder fünfmal, so dass bis zu fünf Gene in einem Konstrukt zusammengeführt werden und so zur Expression in die transgene Pflanze eingebracht werden können. Vorteilhaft wiederholt sich die Abfolge bis zu dreimal. Die Nukleinsäuresequenzen werden zur Expression über die geeignete Schnittstelle beispielsweise im Polylinker hinter den Promotor inseriert. Vorteilhaft hat jede Nukleinsäuresequenz ihren eigenen Promotor und gegebenenfalls ihren eigenen Terminator. Es ist aber auch möglich mehrere Nukleinsäuresequenzen hinter einem Promotor und ggf. vor einem Terminator zu inserieren. Dabei ist die Insertionsstelle bzw. die Abfolge der inserierten Nukleinsäuren in der Expressionskassette nicht von entscheidender Bedeutung, das heißt eine Nukleinsäuresequenz kann an erster oder letzter Stelle in der Kassette inseriert sein, ohne dass dadurch die Expression wesentlich beeinflusst wird. Es können in der Expressionskassette vorteilhaft unterschiedliche Promotoren wie beispielsweise der USP-, LegB4 oder DC3-Promotor und unterschiedliche Terminatoren verwendet werden. Es ist aber auch möglich nur einen Promotortyp in der Kassette zu verwenden. Dies kann jedoch zu unerwünschten Rekombinationsereignissen führen.

Wie oben beschrieben sollte die Transkription der eingebrachten Gene vorteilhaft durch geeignete Terminatoren am 3'-Ende der eingebrachten Biosynthesegene (hinter dem Stoppcodon) abgebrochen werden. Verwendet werden kann hier z.B. der OCS1 Terminator. Wie auch für die Promotoren, so sollten hier für jedes Gen unterschiedliche Terminatorsequenzen verwendet werden.

Das Genkonstrukt kann, wie oben beschrieben, auch weitere Gene umfassen, die in die Organismen eingebracht werden sollen. Es ist möglich und vorteilhaft, in die Wirtsorganismen Regulationsgene, wie Gene für Induktoren, Repressoren oder Enzyme, welche durch ihre Enzymaktivität in die Regulation eines oder mehrerer Gene eines Biosynthesewegs eingreifen, einzubringen und darin zu exprimieren. Diese Gene können heterologen oder homologen Ursprungs sein. Weiterhin können vorteilhaft im Nukleinsäurekonstrukt bzw. Genkonstrukt weitere Biosynthesegene des Fettsäure- oder Lipidstoffwechsels enthalten sein oder aber diese Gene können auf einem weiteren oder mehreren weiteren Nukleinsäurekonstrukten liegen. Vorteilhaft werden als Biosynthesegene des Fettsäure- oder Lipidstoffwechsels ein Gen ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylasen, Lipoxxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) oder deren Kombinationen verwendet. Besonders vorteilhafte Nukleinsäuresequenzen sind Biosynthesegene des Fettsäure- oder Lipidstoffwechsels ausgewählt aus der Gruppe der Δ -4-Desaturase-, Δ -5-Desaturase-, Δ -6-Desaturase-, Δ -8-Desaturase-, Δ -9-Desaturase-, Δ -12-Desaturase-, Δ -5-Elongase-, Δ -6-Elongase- oder Δ -9-Elongase.

Dabei können die vorgenannten Desaturasen in Kombination mit anderen Elongasen und Desaturasen in erfindungsgemäßen Expressionskassetten kloniert werden und zur Transformation von Pflanzen mithilfe von Agrobakterium eingesetzt werden.

5 Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird. Die Expressionskassetten können prinzipiell direkt zum Einbringen in die Pflanze verwendet werden oder aber in einen Vektoren eingebracht werden.

10 Diese vorteilhaften Vektoren, vorzugsweise Expressionsvektoren, enthalten die im Verfahren verwendeten Nukleinsäuren, die für Acyl-CoA:Lysophospholipid-Acyltransferasen codieren, oder ein Nukleinsäurekonstrukt, die die verwendeten Nukleinsäure allein oder in Kombination mit weiteren Biosynthesegenen des Fettsäure- oder Lipidstoffwechsels wie Δ -4-Desaturase-, Δ -5-Desaturase-, Δ -6-Desaturase-, Δ -8-Desaturase-, Δ -9-Desaturase-, Δ -12-Desaturase-, Δ -5-Elongase-, Δ -6-Elongase- und/oder Δ -9-Elongase. Wie hier verwendet, betrifft der Begriff "Vektor" ein Nukleinsäuremolekül, das eine andere Nukleinsäure transportieren kann, an welche es gebunden ist. Ein Vektortyp ist ein "Plasmid", was für eine zirkuläre doppelsträngige DNA-Schleife steht, in die zusätzlichen DNA-Segmente ligiert werden können. Ein weiterer Vektortyp ist ein viraler Vektor, wobei zusätzliche DNA-Segmente in das virale Genom ligiert werden können. Bestimmte Vektoren können in einer Wirtszelle, in die sie eingebracht worden sind, autonom replizieren (z.B. Bakterienvektoren mit bakteriellem Replikationsursprung). Andere Vektoren werden vorteilhaft beim Einbringen in die Wirtszelle in das Genom einer Wirtszelle integriert und dadurch zusammen mit dem Wirtsgenom repliziert. Zudem können bestimmte Vektoren die Expression von Genen, mit denen sie funktionsfähig verbunden sind, steuern. Diese Vektoren werden hier als "Expressionsvektoren" bezeichnet. Gewöhnlich haben Expressionsvektoren, die für DNA-Rekombinationstechniken geeignet sind, die Form von Plasmiden. In der vorliegenden Beschreibung können "Plasmid" und "Vektor" austauschbar verwendet werden, da das Plasmid die am häufigsten verwendete Vektorform ist. Die Erfindung soll jedoch diese anderen Expressionsvektorformen, wie virale Vektoren, die ähnliche Funktionen ausüben, umfassen. Ferner soll der Begriff Vektor auch andere Vektoren, die dem Fachmann bekannt sind, wie Phagen, Viren, wie SV40, CMV, TMV, Transposons, IS-Elemente, Phasmide, Phagemide, Cosmide, lineare oder zirkuläre DNA, umfassen.

Die im Verfahren vorteilhaft verwendeten rekombinanten Expressionsvektoren umfassen die die unten beschriebenen Nukleinsäuren oder das oben beschriebene Genkonstrukt in einer Form, die sich zur Expression der verwendeten Nukleinsäuren in einer Wirtszelle eignen, was bedeutet, dass die rekombinanten Expressionsvektoren eine oder mehrere Regulationssequenzen, ausgewählt auf der Basis der zur Expression zu verwendenden Wirtszellen, die mit der zu exprimierenden Nukleinsäuresequenz

- funktionsfähig verbunden ist, umfasst. In einem rekombinanten Expressionsvektor bedeutet "funktionsfähig verbunden", dass die Nukleotidsequenz von Interesse derart an die Regulationssequenz(en) gebunden ist, dass die Expression der Nukleotidsequenz möglich ist und sie aneinander gebunden sind, so dass beide Sequenzen die vorher-
5. gesagte, der Sequenz zugeschriebene Funktion erfüllen (z.B. in einem In-vitro-Transkriptions-/Translationssystem oder in einer Wirtszelle, wenn der Vektor in die Wirtszelle eingebracht wird). Der Begriff "Regulationssequenz" soll Promotoren, Enhancer und andere Expressionskontrollelemente (z.B. Polyadenylierungssignale) umfassen. Diese Regulationssequenzen sind z.B. beschrieben in Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA
- 10 (1990), oder siehe: Gruber und Crosby, in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, Hrsgb.: Glick und Thompson, Kapitel 7, 89-108, einschließlich der Literaturstellen darin. Regulationssequenzen umfassen solche, welche die konstitutive Expression einer Nukleotidsequenz in vielen Wirtszelltypen steuern, und solche, welche die direkte Expression der Nukleotidsequenz nur in bestimmten Wirtszellen unter bestimmten Bedingungen steuern. Der Fachmann weiß, dass die Gestaltung des Expressionsvektors von Faktoren, wie der Auswahl der zu transformierenden Wirtszelle, dem Ausmaß der Expression des gewünschten Proteins usw., abhängen kann.
- 15
- 20 Die verwendeten rekombinanten Expressionsvektoren können zur Expression von Acyl-CoA:Lysophospholipid-Acyltransferasen, Desaturasen und Elongasen in prokaryotischen oder eukaryotischen Zellen gestaltet sein. Dies ist vorteilhaft, da häufig Zwischenschritte der Vektorkonstruktion der Einfachheit halber in Mikroorganismen durchgeführt werden. Beispielsweise können Acyl-CoA:Lysophospholipid-Acyltransferase-,
- 25 Desaturase- und/oder Elongase-Gene in bakteriellen Zellen, Insektenzellen (unter Verwendung von Baculovirus-Expressionsvektoren), Hefe- und anderen Pilzzellen (siehe Romanos, M.A., et al. (1992) "Foreign gene expression in yeast: a review", Yeast 8:423-488; van den Hondel, C.A.M.J.J., et al. (1991) "Heterologous gene expression in filamentous fungi", in: More Gene Manipulations in Fungi, J.W. Bennet &
- 30 L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego; und van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, Peberdy, J.F., et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge), Algen (Falciatore et al., 1999, Marine Biotechnology. 1, 3:239-251), Ciliaten der Typen: Holotrichia, Peritrichia, Spirotrichia, Suctoria, Tetrahymena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Desaturaseudocohnilembus, Euplotes, Engelmaniella und Stylonychia, insbesondere der Gattung Stylonychia lemnae, mit Vektoren nach einem Transformationsverfahren, wie beschrieben in WO 98/01572, sowie bevorzugt in Zellen vielzelliger Pflanzen (siehe Schmidt, R. und Willmitzer, L. (1988) "High efficiency Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana leaf and cotyledon explants" Plant Cell Rep.:583-586; Plant Molecular Biology and Biotechnology, C Press, Boca Raton, Florida, Kapitel 6/7, S.71-119 (1993); F.F. White, B. Jené et al., Techniques for Gene Transfer, in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-43; Potrykus, Annu. Rev. Plant
- 40

Physiol. Plant Molec. Biol. 42 (1991), 205-225 (und darin zitierte Literaturstellen)) exprimiert werden. Geeignete Wirtszellen werden ferner erörtert in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990). Der rekombinante Expressionsvektor kann alternativ, zum Beispiel unter Verwendung von T7-Promotor-Regulationssequenzen und T7-Polymerase, in vitro transkribiert und translatiert werden.

Die Expression von Proteinen in Prokaryoten erfolgt meist mit Vektoren, die konstitutive oder induzierbare Promotoren enthalten, welche die Expression von Fusions- oder nicht-Fusionsproteinen steuern. Typische Fusions-Expressionsvektoren sind u.a. pGEX (Pharmacia Biotech Inc; Smith, D.B., und Johnson, K.S. (1988) Gene 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRIT5 (Pharmacia, Piscataway, NJ), bei denen Glutathion-S-Transferase (GST), Maltose E-bindendes Protein bzw. Protein A an das rekombinante Zielprotein fusioniert wird.

Beispiele für geeignete induzierbare nicht-Fusions-E. coli-Expressionsvektoren sind u.a. pTrc (Amann et al. (1988) Gene 69:301-315) und pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 60-89). Die Zielgenexpression vom pTrc-Vektor beruht auf der Transkription durch Wirts-RNA-Polymerase von einem Hybrid-trp-lac-Fusionspromotor. Die Zielgenexpression aus dem pET 11d-Vektor beruht auf der Transkription von einem T7-gn10-lac-Fusions-Promotor, die von einer coexprimierten viralen RNA-Polymerase (T7 gn1) vermittelt wird. Diese virale Polymerase wird von den Wirtsstämmen BL21 (DE3) oder HMS174 (DE3) von einem residenten λ -Prophagen bereitgestellt, der ein T7 gn1-Gen unter der Transkriptionskontrolle des lacUV 5-Promotors birgt.

Andere in prokaryotischen Organismen geeignete Vektoren sind dem Fachmann bekannt, diese Vektoren sind beispielsweise in E. coli pLG338, pACYC184, die pBR-Reihe, wie pBR322, die pUC-Reihe, wie pUC18 oder pUC19, die M13mp-Reihe, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1, λ gt11 or pBdCl, in Streptomyces pIJ101, pIJ364, pIJ702 oder pIJ361, in Bacillus pUB110, pC194 oder pBD214, in Corynebacterium pSA77 oder pAJ667.

Bei einer weiteren Ausführungsform ist der Expressionsvektor ein Hefe-Expressionsvektor. Beispiele für Vektoren zur Expression in der Hefe *S. cerevisiae* umfassen pYe-Desaturasec1 (Baldari et al. (1987) Embo J. 6:229-234), pMFa (Kurjan und Herskowitz (1982) Cell 30:933-943), pJRY88 (Schultz et al. (1987) Gene 54:113-123) sowie pYES2 (Invitrogen Corporation, San Diego, CA). Vektoren und Verfahren zur Konstruktion von Vektoren, die sich zur Verwendung in anderen Pilzen, wie den filamentösen Pilzen, eignen, umfassen diejenigen, die eingehend beschrieben sind in: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of fungi, J.F. Peberdy et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge, oder in: More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego].

Weitere geeignete Hefevektoren sind beispielsweise pAG-1, YEp6, YEp13 oder pEMBLye23.

Alternativ können die Acyl-CoA:Lysophospholipid-Acyltransferasen, Desaturasen und/oder Elongasen in Insektenzellen unter Verwendung von Baculovirus-Expressionsvektoren exprimiert werden. Baculovirus-Vektoren, die zur Expression von Proteinen in gezüchteten Insektenzellen (z.B. Sf9-Zellen) verfügbar sind, umfassen die pAc-Reihe (Smith et al. (1983) Mol. Cell Biol. 3:2156-2165) und die pVL-Reihe (Lucklow und Summers (1989) Virology 170:31-39).

Die oben genannten Vektoren bieten nur einen kleinen Überblick über mögliche geeignete Vektoren. Weitere Plasmide sind dem Fachmann bekannt und sind zum Beispiel beschrieben in: Cloning Vectors (Hrsgb. Pouwels, P.H., et al., Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018). Weitere geeignete Expressionssysteme für prokaryotische und eukaryotische Zellen siehe in den Kapiteln 16 und 17 von Sambrook, J., Fritsch, E.F., und Maniatis, T., Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Bei einer weiteren Ausführungsform des Verfahrens können die Acyl-CoA:Lyso-phospholipid-Acyltransferasen, Desaturasen und/oder Elongasen in einzelligen Pflanzenzellen (wie Algen), siehe Falciatore et al., 1999, Marine Biotechnology 1 (3):239-251 und darin zitierte Literaturangaben, und Pflanzenzellen aus höheren Pflanzen (z.B. Spermatophyten, wie Feldfrüchten) exprimiert werden. Beispiele für Pflanzen-Expressionsvektoren umfassen solche, die eingehend beschrieben sind in: Becker, D., Kemper, E., Schell, J., und Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", Plant Mol. Biol. 20:1195-1197; und Bevan, M.W. (1984) "Binary Agrobacterium vectors for plant transformation", Nucl. Acids Res. 12:8711-8721; Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, S. 15-38.

Eine Pflanzen-Expressionskassette enthält vorzugsweise Regulationssequenzen, welche die Genexpression in Pflanzenzellen steuern können und funktionsfähig verbunden sind, so dass jede Sequenz ihre Funktion, wie Termination der Transkription, erfüllen kann, beispielsweise Polyadenylierungssignale. Bevorzugte Polyadenylierungssignale sind diejenigen, die aus Agrobacterium tumefaciens-T-DNA stammen, wie das als Octopinsynthese bekannte Gen 3 des Ti-Plasmids pTiACH5 (Gielen et al., EMBO J. 3 (1984) 835ff.) oder funktionelle Äquivalente davon, aber auch alle anderen in Pflanzen funktionell aktiven Terminatoren sind geeignet.

Da die Pflanzengenexpression sehr oft nicht auf Transkriptionsebenen beschränkt ist, enthält eine Pflanzen-Expressionskassette vorzugsweise andere funktionsfähig verbunden Sequenzen, wie Translationsenhancer, beispielsweise die Overdrive-Sequenz, welche die 5'-untranslatierte Leader-Sequenz aus Tabakmosaikvirus, die das Prote-

in/RNA-Verhältnis erhöht, enthält (Gallie et al., 1987, Nucl. Acids Research 15:8693-8711).

Die Pflanzengenexpression muss wie oben beschrieben funktionsfähig mit einem geeigneten Promotor verbunden sein, der die Genexpression auf rechtzeitige, zell- oder gewebespezifische Weise durchführt. Nutzbare Promotoren sind konstitutive Promotoren (Benfey et al., EMBO J. 8 (1989) 2195-2202), wie diejenigen, die von Pflanzenviren stammen, wie 35S CAMV (Franck et al., Cell 21 (1980) 285-294), 19S CaMV (siehe auch US 5352605 und WO 84/02913) oder Pflanzenpromotoren, wie der in US 4,962,028 beschriebene der kleinen Untereinheit der Rubisco.

Andere bevorzugte Sequenzen für die Verwendung zur funktionsfähigen Verbindung in Pflanzengenexpressions-Kassetten sind Targeting-Sequenzen, die zur Steuerung des Genproduktes in sein entsprechendes Zellkompartiment notwendig sind (siehe eine Übersicht in Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423 und darin zitierte Literaturstellen), beispielsweise in die Vakuole, den Zellkern, alle Arten von Plastiden, wie Amyloplasten, Chloroplasten, Chromoplasten, den extrazellulären Raum, die Mitochondrien, das Endoplasmatische Retikulum, Ölkörper, Peroxisomen und andere Kompartimente von Pflanzenzellen.

Die Pflanzengenexpression lässt sich auch wie oben beschrieben über einen chemisch induzierbaren Promotor erleichtern (siehe eine Übersicht in Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108). Chemisch induzierbare Promotoren eignen sich besonders, wenn gewünscht wird, dass die Genexpression auf zeitspezifische Weise erfolgt. Beispiele für solche Promotoren sind ein Salicylsäure-induzierbarer Promotor (WO 95/19443), ein Tetracyclin-induzierbarer Promotor (Gatz et al. (1992) Plant J. 2, 397-404) und ein Ethanol-induzierbarer Promotor.

Auch Promotoren, die auf biotische oder abiotische Stressbedingungen reagieren, sind geeignete Promotoren, beispielsweise der pathogeninduzierte PRP1-Gen-Promotor (Ward et al., Plant. Mol. Biol. 22 (1993) 361-366), der hitzeinduzierbare hsp80-Promotor aus Tomate (US 5,187,267), der kälteinduzierbare Alphaamylase-Promotor aus Kartoffel (WO 96/12814) oder der durch Wunden induzierbare pinII-Promotor (EP-A-0 375 091).

Es sind insbesondere solche Promotoren bevorzugt, welche die Genexpression in Geweben und Organen herbeiführen, in denen die Fettsäure-, Lipid- und Ölbiosynthese stattfindet, in Samenzellen, wie den Zellen des Endosperms und des sich entwickelnden Embryos. Geeignete Promotoren sind der Napingen-Promotor aus Raps (US 5,608,152), der USP-Promotor aus Vicia faba (Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67), der Oleosin-Promotor aus Arabidopsis (WO 98/45461), der Phaseolin-Promotor aus Phaseolus vulgaris (US 5,504,200), der Bce4-Promotor aus Brassica (WO 91/13980) oder der Legumin-B4-Promotor (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2):233-9) sowie Promotoren, welche die samenspezifische Expression in Monokotyledonen-Pflanzen, wie Mais, Gerste, Weizen, Roggen, Reis usw. herbeiführen. Geeignete beachtenswerte Promotoren sind der lpt2- oder lpt1-Gen-

Promotor aus Gerste (WO 95/15389 und WO 95/23230) oder die in WO 99/16890 beschriebenen (Promotoren aus dem Gersten-Hordein-Gen, dem Reis-Glutelin-Gen, dem Reis-Oryzin-Gen, dem Reis-Prolamin-Gen, dem Weizen-Gliadin-Gen, Weizen-Glutelin-Gen, dem Mais-Zein-Gen, dem Hafer-Glutelin-Gen, dem Sorghum-Kasirin-Gen, dem 5 Roggen-Secalin-Gen).

Insbesondere kann die multiparallele Expression der im Verfahren verwendeten Acyl-CoA:Lysophospholipid-Acyltransferasen allein oder in Kombination mit Desaturasen und/oder Elongasen gewünscht sein. Die Einführung solcher Expressionskassetten kann über eine simultane Transformation mehrerer einzelner Expressionskonstrukte 10 erfolgen oder bevorzugt durch Kombination mehrerer Expressionskassetten auf einem Konstrukt. Auch können mehrere Vektoren mit jeweils mehreren Expressionskassetten transformiert und auf die Wirtszelle übertragen werden.

Ebenfalls, besonders geeignet sind Promotoren, welche die plastidenspezifische Expression herbeiführen, da Plastiden das Kompartiment sind, in dem die Vorläufer sowie 15 einige Endprodukte der Lipidbiosynthese synthetisiert werden. Geeignete Promotoren, wie der virale RNA-Polymerase-Promotor, sind beschrieben in WO 95/16783 und WO 97/06250, und der clpP-Promotor aus Arabidopsis, beschrieben in WO 99/46394.

Vektor-DNA lässt sich in prokaryotische oder eukaryotische Zellen über herkömmliche Transformations- oder Transfektionstechniken einbringen. Die Begriffe "Transformation" und "Transfektion", Konjugation und Transduktion, wie hier verwendet, sollen eine 20 Vielzahl von im Stand der Technik bekannten Verfahren zum Einbringen fremder Nukleinsäure (z.B. DNA) in eine Wirtszelle, einschließlich Calciumphosphat- oder Calciumchlorid-Coprazipitation, DEAE-Dextran-vermittelte Transfektion, Lipofektion, natürliche Kompetenz, chemisch vermittelter Transfer, Elektroporation oder Teilchenbeschuss, umfassen. Geeignete Verfahren zur Transformation oder Transfektion von 25 Wirtszellen, einschließlich Pflanzenzellen, lassen sich finden in Sambrook et al. (Molecular Cloning: A Laboratory Manual., 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) und anderen Labor-Handbüchern, wie Methods in Molecular Biology, 1995, Bd. 44, Agrobacterium protocols, Hrsgb: Gartland und Davey, Humana Press, Totowa, New Jersey. 30

Wirtszellen, die im Prinzip zum Aufnehmen der erfindungsgemäßen Nukleinsäure, des erfindungsgemäßen Genproduktes oder des erfindungsgemäßen Vektors geeignet sind, sind alle prokaryotischen oder eukaryotischen Organismen. Die vorteilhafterweise verwendeten Wirtsorganismen sind Mikroorganismen, wie Pilze oder Hefen oder Pflanzenzellen vorzugsweise Pflanzen oder Teile davon. Pilze, Hefen oder Pflanzen werden 35 vorzugsweise verwendet, besonders bevorzugt Pflanzen, ganz besonders bevorzugt Pflanzen, wie Ölfrüchtpflanzen, die große Mengen an Lipidverbindungen enthalten, wie Raps, Nachtkerze, Hanf, Diestel, Erdnuss, Canola, Lein, Soja, Safflor, Sonnenblume, Borretsch, oder Pflanzen, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa, Buschpflanzen (Kaffee, Kakao, Tee), 40

Salix-Arten, Bäume (Ölpläne, Kokosnuss) sowie ausdauernde Gräser und Futterfeldfrüchte. Besonders bevorzugte erfindungsgemäße Pflanzen sind Ölfruchtpflanzen, wie Soja, Erdnuß, Raps, Canola, Lein, Hanf, Nachtkerze, Sonnenblume, Safflor, Bäume (Ölpläne, Kokosnuß).

- 5 Ein weiterer erfindungsgemäßer Gegenstand sind wie oben beschrieben isolierte Nukleinsäuresequenzen, die für Polypeptide mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codieren; wobei die durch die Nukleinsäuresequenzen codierten Acyl-CoA:Lysophospholipid-Acyltransferasen spezifisch C₁₆-, C₁₈-, C₂₀- oder C₂₂-Fettsäuren mit mindestens einer Doppelbindungen im Fettsäuremolekül umsetzen.
- 10 Vorteilhafte isolierte Nukleinsäuresequenzen sind Sequenzen ausgewählt aus der Gruppe:
- d) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Sequenz,
 - 15 e) Nukleinsäuresequenzen, die sich als Ergebnis des degenerierten genetischen Codes von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 enthaltenden codierenden Sequenz ableiten lassen
 - f) 20 Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 dargestellten Aminosäuresequenz codieren und mindestens 40 % Homologie auf Aminosäureebene mit SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 aufweisen und eine Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität aufweisen.

- 25 Die oben genannte erfindungsgemäßen Nukleinsäuren stammen von Organismen, wie Tieren, Ciliaten, Pilzen, Pflanzen wie Algen oder Dinoflagellaten, die PUFAs synthetisieren können.

- 30 Der Begriff "Nukleinsäure(molekül)", wie hier verwendet, umfasst in einer vorteilhaften Ausführungsform zudem die am 3'- und am 5'-Ende des kodierenden Genbereichs gelegene untranslatierte Sequenz: mindestens 500, bevorzugt 200, besonders bevorzugt 100 Nukleotide der Sequenz stromaufwärts des 5'-Endes des kodierenden Bereichs und mindestens 100, bevorzugt 50, besonders bevorzugt 20 Nukleotide der Sequenz stromabwärts des 3'-Endes des kodierenden Genbereichs. Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure vorliegen. Eine "isolierte" Nukleinsäure hat vorzugsweise keine Sequenzen, welche die Nukleinsäure in der genomischen DNA des Organismus, aus dem die Nukleinsäure stammt, natürlicherweise flankieren (z.B. Sequenzen, die sich an den 5'- und 3'-Enden der Nukleinsäure befinden). Bei verschiedenen Ausführungsformen kann das isolierte Acyl-CoA:Lysophospholipid-Acyltransferasemolekül zum Beispiel weniger als etwa 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5

kb oder 0,1 kb an Nukleotidsequenzen enthalten, die natürlicherweise das Nukleinsäuremolekül in der genomischen DNA der Zelle, aus der die Nukleinsäure stammt flankieren.

Die im Verfahren verwendeten Nukleinsäuremoleküle, z.B. ein Nukleinsäuremolekül mit einer Nukleotidsequenz der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 oder eines Teils davon, kann unter Verwendung molekularbiologischer Standardtechniken und der hier bereitgestellten Sequenzinformation isoliert werden. Auch kann mithilfe von Vergleichsalgorithmen beispielsweise eine homologe Sequenz oder homologe, konservierte Sequenzbereiche auf DNA oder Aminosäureebene identifiziert werden. Diese können als Hybridisierungs-sonde sowie Standard-Hybridisierungstechniken (wie z.B. beschrieben in Sambrook et al., Molecular Cloning: A Laboratory Manual. 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) zur Isolierung weiterer im Verfahren nützlicher Nukleinsäuresequenzen verwendet werden. Überdies lässt sich ein Nukleinsäuremolekül, umfassend eine vollständige Sequenz der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 oder einen Teil davon, durch Polymerasekettenreaktion isolieren, wobei Oligonukleotidprimer, die auf der Basis dieser Sequenz oder von Teilen davon, verwendet werden (z.B. kann ein Nukleinsäuremolekül, umfassend die vollständigen Sequenz oder einen Teil davon, durch Polymerasekettenreaktion unter Verwendung von Oligonukleotidprimern isoliert werden, die auf der Basis dieser gleichen Sequenz erstellt worden sind). Zum Beispiel lässt sich mRNA aus Zellen isolieren (z.B. durch das Guanidiniumthiocyanat-Extraktionsverfahren von Chirgwin et al. (1979) Biochemistry 18:5294-5299) und cDNA mittels Reverse Transkriptase (z.B. Moloney-MLV-Reverse-Transkriptase, erhältlich von Gibco/BRL, Bethesda, MD, oder AMV-Reverse-Transkriptase, erhältlich von Seikagaku America, Inc., St. Petersburg, FL) herstellen. Synthetische Oligonukleotidprimer zur Amplifizierung mittels Polymerasekettenreaktion lassen sich auf der Basis einer der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 gezeigten Sequenzen oder mithilfe der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 dargestellten Aminosäuresequenzen erstellen. Eine erfindungsgemäße Nukleinsäure kann unter Verwendung von cDNA oder alternativ von genomischer DNA als Matrize und geeigneten Oligonukleotidprimern gemäß Standard-PCR-Amplifikationstechniken amplifiziert werden. Die so amplifizierte Nukleinsäure kann in einen geeigneten Vektor kloniert werden und mittels DNA-Sequenzanalyse charakterisiert werden. Oligonukleotide, die einer Desaturase-Nukleotidsequenz entsprechen, können durch Standard-Syntheseverfahren, beispielsweise mit einem automatischen DNA-Synthesegerät, hergestellt werden.

Homologe der verwendeten Acyl-CoA:Lysophospholipid-Acyltransferase-Nukleinsäuresequenzen mit der Sequenz SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 bedeutet beispielsweise allelische Varianten mit mindestens etwa 40 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 %, stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 % oder 90 bis 95 % und noch stärker bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr Homologie zu einer in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 gezeigten Nukleotidsequenzen oder ihren

- Homologen, Derivaten oder Analoga oder Teilen davon. Weiterhin sind isolierte Nukleinsäuremoleküle einer Nukleotidsequenz, die an eine der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 gezeigten Nukleotidsequenzen oder einen Teil davon hybridisieren, z.B. unter stringenten Bedingungen hybridisiert. Allelische Varianten umfassen insbesondere funktionelle Varianten, die sich durch Deletion, Insertion oder Substitution von Nukleotiden aus/in der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Sequenz erhalten lassen, wobei aber die Absicht ist, dass die Enzymaktivität der davon herrührenden synthetisierten Proteine für die Insertion eines oder mehrerer Gene vorteilhafterweise beibehalten wird. Proteine, die noch die enzymatische Aktivität der Acyl-CoA:Lysophospholipid-Acyltransferase besitzen, das heißt deren Aktivität im wesentlichen nicht reduziert ist, bedeutet Proteine mit mindestens 10 %, vorzugsweise 20 %, besonders bevorzugt 30 %, ganz besonders bevorzugt 40 % der ursprünglichen Enzymaktivität, verglichen mit dem durch SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 kodierten Protein.

- Homologen der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 bedeuten beispielsweise auch bakterielle, Pilz- und Pflanzenhomologen, verkürzte Sequenzen, einzelsträngige DNA oder RNA der kodierenden und nicht-kodierenden DNA-Sequenz.
- Homologen der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 bedeutet auch Derivate, wie beispielsweise Promotorvarianten. Die Promotoren stromaufwärts der angegebenen Nukleotidsequenzen können durch einen oder mehrere Nukleotidaustausche, durch Insertion(en) und/oder Deletion(en) modifiziert werden, ohne dass jedoch die Funktionalität oder Aktivität der Promotoren gestört wird. Es ist weiterhin möglich, dass die Aktivität der Promotoren durch Modifikation ihrer Sequenz erhöht ist oder dass sie vollständig durch aktivere Promotoren, sogar aus heterologen Organismen, ersetzt werden.

- Die vorgenannten Nukleinsäuren und Proteinmoleküle mit Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität, die am Stoffwechsel von Lipiden und Fettsäuren, PUFA-Cofaktoren und Enzymen oder am Transport lipophiler Verbindungen über Membranen beteiligt sind, werden im erfindungsgemäßen Verfahren zur Modulation der Produktion von PUFAs in transgenen Organismen vorteilhaft in Pflanzen, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Sojabohne, Erdnuss, Baumwolle, Linum Arten wie Öl- oder Faserlein, Brassica-Arten, wie Raps, Canola und Rübsen, Pfeffer, Sonnenblume, Borretsch, Nachtkerze und Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Maniok, Alfalfa, Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten, Bäume (Ölpalme, Kokosnuss) und ausdauernden Gräsern und Futterfeldfrüchten, entweder direkt (z.B. wenn die Überexpression oder Optimierung eines Fettsäurebiosynthese-Proteins einen direkten Einfluss auf die Ausbeute, Produktion und/oder Effizienz der Produktion der Fettsäure aus modifizierten Organismen hat) verwendet und/oder können eine indirekt Auswirkung haben, die dennoch zu einer Steigerung der Ausbeute, Produktion und/oder Effizienz der Produktion der PU-

FAs oder einer Abnahme unerwünschter Verbindungen führt (z.B. wenn die Modulation des Stoffwechsels von Lipiden und Fettsäuren, Cofaktoren und Enzymen zu Veränderungen der Ausbeute, Produktion und/oder Effizienz der Produktion oder der Zusammensetzung der gewünschten Verbindungen innerhalb der Zellen führt, was wiederum die Produktion einer oder mehrerer Fettsäuren beeinflussen kann).

Die Kombination verschiedener Vorläufermoleküle und Biosyntheseenzyme führt zur Herstellung verschiedener Fettsäuremoleküle, was eine entscheidende Auswirkung auf die Zusammensetzung der Lipide hat. Da mehrfach ungesättigte Fettsäuren (= PUFAs) nicht nur einfach in Triacylglycerin sondern auch in Membranlipide eingebaut werden.

- 10 Die Lipidsynthese lässt sich in zwei Abschnitte unterteilen: die Synthese von Fettsäuren und ihre Bindung an sn-Glycerin-3-Phosphat sowie die Addition oder Modifikation einer polaren Kopfgruppe. Übliche Lipide, die in Membranen verwendet werden, umfassen Phospholipide, Glycolipide, Sphingolipide und Phosphoglyceride. Die Fettsäuresynthese beginnt mit der Umwandlung von Acetyl-CoA in Malonyl-CoA durch die
- 15 Acetyl-CoA-Carboxylase oder in Acetyl-ACP durch die Acetyltransacylase. Nach einer Kondensationsreaktion bilden diese beiden Produktmoleküle zusammen Acetoacetyl-ACP, das über eine Reihe von Kondensations-, Reduktions- und Dehydratisierungsreaktionen umgewandelt wird, so dass ein gesättigtes Fettsäuremolekül mit der gewünschten Kettenlänge erhalten wird. Die Produktion der ungesättigten Fettsäuren aus
- 20 diesen Molekülen wird durch spezifische Desaturasen katalysiert, und zwar entweder aerob mittels molekularem Sauerstoff oder anaerob (bezüglich der Fettsäuresynthese in Mikroorganismen siehe F.C. Neidhardt et al. (1996) *E. coli* und *Salmonella*. ASM Press: Washington, D.C., S. 612-636 und darin enthaltene Literaturstellen; Lengeler et al. (Hrsgb.) (1999) *Biology of Prokaryotes*. Thieme: Stuttgart, New York, und die enthaltene Literaturstellen, sowie Magnuson, K., et al. (1993) *Microbiological Reviews* 57:522-542 und die enthaltenen Literaturstellen). Die so hergestellten an
- 25 Phospholipide gebundenen Fettsäuren müssen anschließend wieder für die weitere Elongationen aus den Phospholipiden in den FettsäureCoA-Ester-Pool überführt werden. Dies ermöglichen die erfindungsgemäßen Acyl-CoA:Lysophospholipid-
- 30 Acyltransferasen. Weiterhin können diese Enzyme die elongierten Fettsäuren wieder von den CoA-Estern auf die Phospholipide übertragen. Diese Reaktionsabfolge kann gegebenenfalls mehrfach durchlaufen werden (siehe Figur 10).

- Vorläufer für die PUFA-Biosynthese sind beispielsweise Ölsäure, Linol- und Linolensäure. Diese C_{18} -Kohlenstoff-Fettsäuren müssen auf C_{20} und C_{22} verlängert werden,
- 35 damit Fettsäuren vom Eicosa- und Docosa-Kettentyp erhalten werden. Mithilfe der im Verfahren verwendeten Acyl-CoA:Lysophospholipid-Acyltransferasen, vorteilhaft in Kombination mit Desaturasen wie der Δ -4-, Δ -5-, Δ -6- und Δ -8-Desaturasen und/oder der Δ -5-, Δ -6-, Δ -9-Elongase können Arachidonsäure, Eicosapentaensäure, Docosapentaensäure oder Docosahexaensäure sowie verschiedene andere langkettige PU-
- 40 FAs erhalten, extrahiert und für verschiedene Zwecke bei Nahrungsmittel-, Futter-, Kosmetik- oder pharmazeutischen Anwendungen verwendet werden. Mit den genannten Enzymen können vorzugsweise C_{18} -, C_{20} -, und/oder C_{22} -Fettsäuren mit mindestens

- zwei vorteilhaft mindestens drei, vier, fünf oder sechs Doppelbindungen im Fettsäuremolekül, vorzugsweise zu C₂₀-, und/oder C₂₂-Fettsäuren mit vorteilhaft drei, vier oder fünf Doppelbindungen im Fettsäuremolekül hergestellt werden. Die Desaturierung kann vor oder nach Elongation der entsprechenden Fettsäure erfolgen. Daher führen die
- 5,3 Produkte der Desaturaseaktivitäten und der möglichen weiteren Desaturierung und Elongation zu bevorzugten PUFAs mit höherem Desaturierungsgrad, einschließlich einer weiteren Elongation von C₂₀ zu C₂₂-Fettsäuren, zu Fettsäuren wie γ -Linolensäure, Dihomo- γ -linolensäure, Arachidonsäure, Stearidonsäure, Eicosatetraensäure oder Eicosapentaensäure. Substrate Substrat der Acyl-CoA:Lysophospholipid-
- 10 Acyltransferasen im erfindungsgemäßen Verfahren sind C₁₆-, C₁₈-, C₂₀- oder C₂₂-Fettsäuren wie zum Beispiel Palmitinsäure, Palmitoleinsäure, Linolsäure, γ -Linolensäure, α -Linolensäure, Dihomo- γ -linolensäure, Eicosatetraensäure oder Stearidonsäure. Bevorzugte Substrate sind Linolsäure, γ -Linolensäure und/oder α -Linolensäure, Dihomo- γ -linolensäure bzw. Arachidonsäure, Eicosatetraensäure oder
- 15 Eicosapentaensäure. Die C₁₈-, C₂₀- oder C₂₂-Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure fallen im erfindungsgemäßen Verfahren in Form der freien Fettsäure oder in Form ihrer Ester beispielsweise in Form ihrer Glyceride an.

- Unter dem Begriff "Glycerid" wird ein mit ein, zwei oder drei Carbonsäureresten verestertes Glycerin verstanden (Mono-, Di- oder Triglycerid). Unter "Glycerid" wird auch
- 20 ein Gemisch an verschiedenen Glyceriden verstanden. Das Glycerid oder das Glyceridgemisch kann weitere Zusätze, z.B. freie Fettsäuren, Antioxidantien, Proteine, Kohlenhydrate, Vitamine und/oder andere Substanzen enthalten.

- Unter einem "Glycerid" im Sinne des erfindungsgemäßen Verfahrens werden ferner vom Glycerin abgeleitete Derivate verstanden. Dazu zählen neben den oben beschriebenen Fettsäureglyceriden auch Glycerophospholipide und Glyceroglycolipide. Bevorzugt seien hier die Glycerophospholipide wie Lecithin (Phosphatidylcholin), Cardiolipin, Phosphatidylglycerin, Phosphatidylserin und Alkylacylglycerophospholipide beispielhaft
- 25 genannt.

- Ferner müssen Fettsäuren anschließend an verschiedene Modifikationsorte transportiert und in das Triacylglycerin-Speicherlipid eingebaut werden. Ein weiterer wichtiger Schritt bei der Lipidsynthese ist der Transfer von Fettsäuren auf die polaren Kopfgruppen, beispielsweise durch Glycerin-Fettsäure-Acyltransferase (siehe Frentzen, 1998, Lipid, 100(4-5):161-166).
- 30

- Veröffentlichungen über die Pflanzen-Fettsäurebiosynthese, Desaturierung, den Lipidstoffwechsel und Membrantransport von fetthaltigen Verbindungen, die Betaoxidation, Fettsäuremodifikation und Cofaktoren, Triacylglycerin-Speicherung und -Assemblerung einschließlich der Literaturstellen darin siehe in den folgenden Artikeln: Kinney, 1997, Genetic Engineering, Hrsgb.: JK Setlow, 19:149-166; Ohlrogge und Browse, 1995, Plant Cell 7:957-970; Shanklin und Cahoon, 1998, Annu. Rev. Plant
- 35 Physiol. Plant Mol. Biol. 49:611-641; Voelker, 1996, Genetic Engineering, Hrsgb.: JK Setlow, 18:111-13; Gerhardt, 1992, Prog. Lipid R. 31:397-417; Gühnemann-Schäfer &
- 40

Kindl, 1995, Biochim. Biophys Acta 1256:181-186; Kunau et al., 1995, Prog. Lipid Res. 34:267-342; Stymne et al., 1993, in: Biochemistry and Molecular Biology of Membrane and Storage Lipids of Plants, Hrsgb.: Murata und Somerville, Rockville, American Society of Plant Physiologists, 150-158, Murphy & Ross 1998, Plant Journal. 13(1):1-16.

- 5 Die im Verfahren hergestellten PUFAs, umfassen eine Gruppe von Molekülen, die höhere Tiere nicht mehr synthetisieren können und somit aufnehmen müssen oder die höhere Tiere nicht mehr ausreichend selbst herstellen können und somit zusätzlich aufnehmen müssen, obwohl sie leicht von anderen Organismen, wie Bakterien, synthetisiert werden, beispielsweise können Katzen Arachidonsäure nicht mehr synthetisieren.
- 10
- Der Begriff " Acyl-CoA:Lysophospholipid-Acyltransferasen " im Sinne der Erfindung umfasst Proteine, die am Transfer der an Phospholipide gebundenen Fettsäuren in den CoA-Ester-Pool und vice versa teilnehmen, sowie ihre Homologen, Derivaten oder Analoga. Unter Phospholipiden im Sinne der Erfindung sind zu verstehen Phosphatidylcholin, Phosphatidylethanolamin, Phosphatidylserin, Phosphatidylglycerin, und/oder Phosphatidylinositol vorteilhafterweise Phosphatidylcholin. Die Begriffe Acyl-CoA:Lysophospholipid-Acyltransferase-Nukleinsäuresequenz(en) umfassen Nukleinsäuresequenzen, die eine Acyl-CoA:Lysophospholipid-Acyltransferase kodieren und bei denen ein Teil eine kodierende Region und ebenfalls entsprechende 5'- und 3'-untranslatierte Sequenzbereiche sein können. Die Begriffe Produktion oder Produktivität sind im Fachgebiet bekannt und beinhalten die Konzentration des Fermentationsproduktes (Verbindungen der Formel I), das in einer bestimmten Zeitspanne und einem bestimmten Fermentationsvolumen gebildet wird (z.B. kg Produkt pro Stunde pro Liter). Der Begriff Effizienz der Produktion umfasst die Zeit, die zur Erzielung einer bestimmten Produktionsmenge nötig ist (z.B. wie lange die Zelle zur Aufrichtung einer bestimmten Durchsatzrate einer Feinchemikalie benötigt). Der Begriff Ausbeute oder Produkt/Kohlenstoff-Ausbeute ist im Fachgebiet bekannt und umfasst die Effizienz der Umwandlung der Kohlenstoffquelle in das Produkt (d.h. die Feinchemikalie). Dies wird gewöhnlich beispielsweise ausgedrückt als kg Produkt pro kg Kohlenstoffquelle. Durch Erhöhen der Ausbeute oder Produktion der Verbindung wird die Menge der gewonnenen Moleküle oder der geeigneten gewonnenen Moleküle dieser Verbindung in einer bestimmten Kulturmenge über einen festgelegten Zeitraum erhöht. Die Begriffe Biosynthese oder Biosyntheseweg sind im Fachgebiet bekannt und umfassen die Synthese einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle aus Zwischenverbindungen, beispielsweise in einem Mehrschritt- und stark regulierten Prozess. Die Begriffe Abbau oder Abbauweg sind im Fachgebiet bekannt und umfassen die Spaltung einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle in Abbauprodukte (allgemeiner gesagt, kleinere oder weniger komplexe Moleküle) beispielsweise in einem Mehrschritt- und stark regulierten Prozess. Der Begriff Stoffwechsel ist im Fachgebiet bekannt und umfasst die Gesamtheit der biochemischen Reaktionen, die in einem Organismus stattfinden. Der Stoffwechsel einer bestimmten Verbindung (z.B. der Stoffwechsel einer Fettsäure) umfasst dann die Ge-
- 15
- 20
- 25
- 30
- 35
- 40

samtheit der Biosynthese-, Modifikations- und Abbauege dieser Verbindung in der Zelle, die diese Verbindung betreffen.

Bei einer weiteren Ausführungsform kodieren Derivate des erfindungsgemäßen Nukleinsäuremoleküls wieder gegeben in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 Proteine mit mindestens 40 %, vorteilhaft etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder mehr Homologie (= Identität) zu einer vollständigen Aminosäuresequenz der SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8. Die Homologie wurde über den gesamten Aminosäure- bzw. Nukleinsäuresequenzbereich berechnet. Für die Sequenzvergleiche wurde das Programm PileUp verwendet (J. Mol. Evolution., 25, 351-360, 1987, Higgins et al., CABIOS, 5 1989: 151-153) oder die Programme Gap und BestFit [Needleman and Wunsch (J. Mol. Biol. 48; 443-453 (1970) und Smith and Waterman (Adv. Appl. Math. 2; 482-489 (1981))], die im GCG Software-Packet [Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711 (1991)] enthalten sind. Die oben in Prozent angegebenen Sequenzhomologiewerte wurden mit dem Programm BestFit über den gesamten Sequenzbereich mit folgenden Einstellungen ermittelt: Gap Weight: 8, Length Weight: 2.

Die Erfindung umfasst zudem Nukleinsäuremoleküle, die sich von einer der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 gezeigten Nukleotidsequenzen (und Teilen davon) aufgrund des degenerierten genetischen Codes unterscheiden und somit die gleiche Acyl-CoA:Lysophospholipid-Acyltransferase kodieren wie diejenige, die von den in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 gezeigten Nukleotidsequenzen kodiert wird.

Zusätzlich zu den in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 gezeigten Acyl-CoA:Lysophospholipid-Acyltransferase-Nukleotidsequenzen erkennt der Fachmann, dass DNA-Sequenzpolymorphismen, die zu Änderungen in den Aminosäuresequenzen der Acyl-CoA:Lysophospholipid-Acyltransferasen führen, innerhalb einer Population existieren können. Diese genetischen Polymorphismen im Acyl-CoA:Lysophospholipid-Acyltransferase-Gen können zwischen Individuen innerhalb einer Population aufgrund von natürlicher Variation existieren. Diese natürlichen Varianten bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz des Acyl-CoA:Lysophospholipid-Acyltransferase-Gens. Sämtliche und alle dieser Nukleotidvariationen und daraus resultierende Aminosäurepolymorphismen in der Acyl-CoA:Lysophospholipid-Acyltransferase, die das Ergebnis natürlicher Variation sind und die funktionelle Aktivität von Acyl-CoA:Lysophospholipid-Acyltransferasen nicht verändern, sollen im Umfang der Erfindung enthalten sein.

Für das erfindungsgemäße Verfahren vorteilhafte Nukleinsäuremoleküle können auf der Grundlage ihrer Homologie zu den hier offenbarten Acyl-CoA:Lysophospholipid-Acyltransferase-Nukleinsäuren unter Verwendung der Sequenzen oder eines Teils davon als Hybridisierungsprobe gemäß Standard-

Hybridisierungstechniken unter stringenten Hybridisierungsbedingungen isoliert werden. Dabei können beispielsweise isolierte Nukleinsäuremoleküle verwendet werden, die mindestens 15 Nukleotide lang sind und unter stringenten Bedingungen mit dem Nukleinsäuremolekülen, die eine Nukleotidsequenz der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 umfassen, hybridisieren. Es können auch Nukleinsäuren mindestens 25, 50, 100, 250 oder mehr Nukleotide verwendet werden. Der Begriff "hybridisiert unter stringenten Bedingungen", wie hier verwendet, soll Hybridisierungs- und Waschbedingungen beschreiben, unter denen Nukleotidsequenzen, die mindestens 60 % homolog zueinander sind, gewöhnlich aneinander hybridisiert bleiben. Die Bedingungen sind vorzugsweise derart, dass Sequenzen, die mindestens etwa 65 %, stärker bevorzugt mindestens etwa 70 % und noch stärker bevorzugt mindestens etwa 75 % oder stärker zueinander homolog sind, gewöhnlich aneinander hybridisiert bleiben. Diese stringenten Bedingungen sind dem Fachmann bekannt und lassen sich in Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989), 6.3.1-6.3.6., finden. Ein bevorzugtes, nicht einschränkendes Beispiel für stringente Hybridisierungsbedingungen sind Hybridisierungen in 6 x Natriumchlorid/Natriumcitrat (sodium chloride/sodium citrate = SSC) bei etwa 45°C, gefolgt von einem oder mehreren Waschschritten in 0,2 x SSC, 0,1 % SDS bei 50 bis 65°C. Dem Fachmann ist bekannt, dass diese Hybridisierungsbedingungen sich je nach dem Typ der Nukleinsäure und, wenn beispielsweise organische Lösungsmittel vorliegen, hinsichtlich der Temperatur und der Konzentration des Puffers unterscheiden. Die Temperatur unterscheidet sich beispielsweise unter "Standard-Hybridisierungsbedingungen" je nach dem Typ der Nukleinsäure zwischen 42°C und 58°C in wässrigem Puffer mit einer Konzentration von 0,1 bis 5 x SSC (pH 7,2). Falls organisches Lösungsmittel im obengenannten Puffer vorliegt, zum Beispiel 50 % Formamid, ist die Temperatur unter Standardbedingungen etwa 42°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:DNA-Hybride zum Beispiel 0,1 x SSC und 20°C bis 45°C, vorzugsweise zwischen 30°C und 45°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:RNA-Hybride zum Beispiel 0,1 x SSC und 30°C bis 55°C, vorzugsweise zwischen 45°C und 55°C. Die vorstehend genannten Hybridisierungstemperaturen sind beispielsweise für eine Nukleinsäure mit etwa 100 bp (= Basenpaare) Länge und einem G + C-Gehalt von 50 % in Abwesenheit von Formamid bestimmt. Der Fachmann weiß, wie die erforderlichen Hybridisierungsbedingungen anhand von Lehrbüchern, wie dem vorstehend erwähnten oder aus den folgenden Lehrbüchern Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989; Hames und Higgins (Hrsgb.) 1985, "Nucleic Acids Hybridization: A Practical Approach", IRL Press at Oxford University Press, Oxford; Brown (Hrsgb.) 1991, "Essential Molecular Biology: A Practical Approach", IRL Press at Oxford University Press, Oxford, bestimmt werden können.

Zur Bestimmung der prozentualen Homologie (= Identität) von zwei Aminosäuresequenzen (z.B. einer der Sequenzen der SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8) oder von zwei Nukleinsäuren (z.B. SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7) werden die Sequenzen zum Zweck des optimalen Vergleichs untereinander geschrieben (z.B. können Lücken in die Sequenz eines Proteins oder einer Nukleinsäure eingefügt werden, um ein optimales Alignment mit dem

anderen Protein oder der anderen Nukleinsäure zu erzeugen). Die Aminosäurereste oder Nukleotide an den entsprechenden Aminosäurepositionen oder Nukleotidpositionen werden dann verglichen. Wenn eine Position in einer Sequenz durch den gleichen Aminosäurerest oder das gleiche Nukleotid wie die entsprechende Stelle in der anderen Sequenz belegt wird, dann sind die Moleküle an dieser Position homolog (d.h. Aminosäure- oder Nukleinsäure-"Homologie", wie hier verwendet, entspricht Aminosäure- oder Nukleinsäure-"Identität"). Die prozentuale Homologie zwischen den beiden Sequenzen ist eine Funktion der Anzahl an identischen Positionen, die den Sequenzen gemeinsam sind (d.h. % Homologie = Anzahl der identischen Positionen/Gesamtanzahl der Positionen x 100). Die Begriffe Homologie und Identität sind damit als Synonym anzusehen. Die verwendeten Programme bzw. Algorithmen sind oben beschrieben.

Ein isoliertes Nukleinsäuremolekül, das eine Acyl-CoA:Lysophospholipid-Acyltransferase kodiert, die zu einer Proteinsequenz der SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 homolog ist, kann durch Einbringen einer oder mehrerer Nukleotidsubstitutionen, -additionen oder -deletionen in eine Nukleotidsequenz der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 erzeugt werden, so dass eine oder mehrere Aminosäuresubstitutionen, -additionen oder -deletionen in das kodierte Protein eingebracht werden. Mutationen können in eine der Sequenzen der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 durch Standardtechniken, wie stellenspezifische Mutagenese und PCR-vermittelte Mutagenese, eingebracht werden. Vorzugsweise werden konservative Aminosäuresubstitutionen an einer oder mehreren der vorhergesagten nicht-essentiellen Aminosäureresten hergestellt. Bei einer "konservativen Aminosäuresubstitution" wird der Aminosäurerest gegen einen Aminosäurerest mit einer ähnlichen Seitenkette ausgetauscht. Im Fachgebiet sind Familien von Aminosäureresten mit ähnlichen Seitenketten definiert worden. Diese Familien umfassen Aminosäuren mit basischen Seitenketten (z.B. Lysin, Arginin, Histidin), sauren Seitenketten (z.B. Asparaginsäure, Glutaminsäure), ungeladenen polaren Seitenketten (z.B. Glycin, Asparagin, Glutamin, Serin, Threonin, Tyrosin, Cystein), unpolaren Seitenketten, (z.B. Alanin, Valin, Leucin, Isoleucin, Prolin, Phenylalanin, Methionin, Tryptophan), beta-verzweigten Seitenketten (z.B. Threonin, Valin, Isoleucin) und aromatischen Seitenketten (z.B. Tyrosin, Phenylalanin, Tryptophan, Histidin). Ein vorhergesagter nicht-essentieller Aminosäurerest in einer Acyl-CoA:Lysophospholipid-Acyltransferase wird somit vorzugsweise durch einen anderen Aminosäurerest aus der gleichen Seitenkettenfamilie ausgetauscht. Alternativ können bei einer anderen Ausführungsform die Mutationen zufallsgemäß über die gesamte oder einen Teil der Acyl-CoA:Lysophospholipid-Acyltransferase-kodierenden Sequenz eingebracht werden, z.B. durch Sättigungsmutagenese, und die resultierenden Mutanten können nach der hier beschriebenen Acyl-CoA:Lysophospholipid-Acyltransferase-Aktivität durchmustert werden, um Mutanten zu identifizieren, die die Acyl-CoA:Lysophospholipid-Acyltransferase-Aktivität beibehalten haben. Nach der Mutagenese einer der Sequenzen der SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 kann das kodierte Protein rekombinant exprimiert werden, und die Aktivität des Proteins kann z.B. unter Verwendung der hier beschriebenen Tests bestimmt werden.

Diese Erfindung wird durch die nachstehenden Beispiele weiter veranschaulicht, die nicht als beschränkend aufgefasst werden sollten. Der Inhalt sämtlicher in dieser Patentanmeldung zitierten Literaturstellen, Patentanmeldungen, Patente und veröffentlichten Patentanmeldungen ist hier durch Bezugnahme aufgenommen.

5 Beispiele

Beispiel 1: Allgemeine Verfahren

a) Allgemeine Klonierungsverfahren:

10 Klonierungsverfahren, wie beispielsweise Restriktionsspaltungen, Agarosegelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrocellulose- und Nylonmembranen, Verbindung von DNA-Fragmenten, Transformation von *Escherichia coli*- und Hefe-Zellen, Anzucht von Bakterien und Sequenzanalyse rekombinanter DNA, wurden durchgeführt wie beschrieben in Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) oder Kaiser, Michaelis und Mitchell (1994) "Methods in Yeast Genetics" (Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3).

b) Chemikalien

20 Die verwendeten Chemikalien wurden, wenn im Text nicht anders angegeben, in p. A.-Qualität von den Firmen Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deisenhofen) bezogen. Lösungen wurden unter Verwendung von reinem pyrogenfreiem Wasser, im nachstehenden Text als H₂O bezeichnet, aus einer Milli-Q-Wassersystem-Wasserreinigungsanlage (Millipore, Eschborn) hergestellt. Restriktionsendonukleasen, DNA-modifizierende Enzyme und molekularbiologische Kits wurden bezogen von den Firmen AGS (Heidelberg), Amersham (Braunschweig), Biometra (Göttingen), Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) und Stratagene (Amsterdam, Niederlande). Wenn nicht anders angegeben, wurden sie nach den Anweisungen des Herstellers verwendet.

c) Klonierung und Expression von Desaturasen und Elongasen

30 Der *Escherichia coli*-Stamm XL1 Blue MRF' kan (Stratagene) wurde zur Subklonierung der Δ -6-Desaturase aus *Physcomitrella patens* verwendet. Für die funktionelle Expression dieses Gens wurde der *Saccharomyces cerevisiae*-Stamm INVSc 1 (Invitrogen Co.) verwendet. *E. coli* wurde in Luria-Bertani-Brühe (LB, Duchefa, Haarlem, Niederlande) bei 37°C kultiviert. Wenn nötig, wurde Ampicillin (100 mg/Liter) zugegeben, und 1,5 % Agar (Gew./Vol.) wurde für feste LB-Medien hinzugefügt. *S. cerevisiae* wurde bei 30°C entweder in YPG-Medium oder in komplettem Minimalmedium ohne Uracil (CMdum; siehe in: Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Albright, L.B., Coen, D.M., und Varki, A. (1995) Current

Protocols in Molecular Biology, John Wiley & Sons, New York) mit entweder 2 % (Gew./Vol.) Raffinose oder Glucose kultiviert. Für feste Medien wurden 2 % (Gew./Vol.) Bacto™-Agar (Difco) hinzugefügt. Die zur Klonierung und Expression verwendeten Plasmide sind pUC18 (Pharmacia) und pYES2 (Invitrogen Co.).

5 d) Klonierung und Expression PUFA-spezifischer Desaturasen und Elongaen

Für die Expression in Pflanzen wurden cDNA Klone aus SEQ ID NO: 9, 11 oder 13 so modifiziert, dass lediglich die Codierregion mittels Polymerase Kettenreaktion unter Zuhilfenahme zweier Oligonukleotide amplifiziert werden. Dabei wurde darauf geachtet, dass eine Konsensussequenz vor dem Startcodon zur effizienten Translation eingehalten wurde. Entweder wurde hierzu die Basenfolge ATA oder AAA gewählt und vor
10 das ATG in die Sequenz eingefügt [Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-2929]. Vor diesem Konsensustriplett wurde zusätzlich eine Restriktions-schnittstelle eingeführt, die kompatibel sein muss zur Schnittstelle des Zielvektors,
15 in den das Fragment kloniert werden soll und mit dessen Hilfe die Genexpression in Mikroorganismen oder Pflanzen erfolgen soll.

Die PCR-Reaktion wurde mit Plasmid-DNA als Matrize in einem Thermocycler (Bio-metra) mit der Pfu-DNA-(Stratagene) Polymerase und dem folgenden Temperaturprogramm durchgeführt: 3 min bei 96°C, gefolgt von 30 Zyklen mit 30 s bei 96°C, 30 s bei
20 55°C und 2 min bei 72°C, 1 Zyklus mit 10 min bei 72°C und Stop bei 4°C. Die Anlagerungstemperatur wurde je nach gewählten Oligonukleotiden variiert. Pro Kilobasenpaare DNA ist von einer Syntheszeit von etwa einer Minute auszugehen. Weitere Parameter, die Einfluss auf die PCR haben wie z.B. Mg-Ionen, Salz, DNA Polymerase etc., sind dem Fachmann auf dem Gebiet geläufig und können nach Bedarf variiert
25 werden.

Die korrekte Größe des amplifizierten DNA-Fragments wurde mittels Agarose-TBE-Gelelektrophorese bestätigt. Die amplifizierte DNA wurde aus dem Gel mit dem QIA-quick-Gelextraktionskit (QIAGEN) extrahiert und in die SmaI-Restriktionsstelle des dephosphorylierten Vektors pUC18 unter Verwendung des Sure Clone Ligations Kit
30 (Pharmacia) ligiert, wobei die pUC-Derivate erhalten wurden. Nach der Transformation von E. coli XL1 Blue MRF' kan wurde eine DNA-Minipräparation [Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid mini-preparation. BioTechniques 4, 310-313] an ampicillinresistenten Transformanden durchgeführt, und positive Klone mittels BamHI-Restriktionsanalyse identifiziert.
35 Die Sequenz des klonierten PCR-Produktes wurde mittels Resequenzierung unter Verwendung des ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt) bestätigt.

e) Transformation von *Agrobacterium*

Die *Agrobacterium*-vermittelte Pflanzentransformation wurde, wenn nicht anders beschrieben, wie von Deblaere et al. (1984, Nucl. Acids Res. 13, 4777-4788) mit Hilfe eines *Agrobacterium tumefaciens*-Stamms durchgeführt.

5 f) Pflanzentransformation

Die *Agrobacterium*-vermittelte Pflanzentransformation wurde, wenn nicht anders beschrieben, unter Verwendung von Standard-Transformations- und Regenerationstechniken durchgeführt (Gelvin, Stanton B., Schilperoort, Robert A., Plant Molecular Biology Manual, 2. Aufl., Dordrecht: Kluwer Academic Publ., 1995, in Sect., Ringbuc Zentrale
10 Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

Nach diesen kann beispielsweise Raps mittels Kotyledonen- oder Hypokotyltransformation transformiert werden (Moloney et al., Plant Cell 8 (1989) 238-242; De Block et al., Plant Physiol. 91 (1989) 694-701). Die Verwendung von Antibiotika für die *Agrobacterium*- und Pflanzenselektion hängt von dem für die Transformation verwendeten binären Vektor und *Agrobacterium*-Stamm ab. Die Rapsselektion wird dabei gewöhnlich unter Verwendung von Kanamycin als selektierbarem Pflanzenmarker durchgeführt.

20 Die Transformation von Soja kann unter Verwendung von beispielsweise einer in EP-A-0 0424 047 (Pioneer Hi-Bred International) oder in EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo) beschriebenen Technik durchgeführt werden.

25 Die Pflanzentransformation unter Verwendung von Teilchenbeschuss, Polyethylenglycol-vermittelter DNA-Aufnahme oder über die Siliziumcarbonatfaser-Technik ist beispielsweise beschrieben von Freeling und Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

30 Der *Agrobacterium*-vermittelte Gentransfer in Lein (*Linum usitatissimum*) wurde, wenn nicht anders beschrieben, wie bei Mlynarova et al. [(1994) Plant Cell Report 13:282-285] beschriebenen Technik durchgeführt.

g) Plasmide für die Pflanzentransformation

Zur Pflanzentransformation wurden binäre Vektoren auf Basis der Vektoren pBinAR (Höfgen und Willmitzer, Plant Science 66 (1990) 221-230) oder pGPTV (Becker et al 1992, Plant Mol. Biol. 20:1195-1197) verwendet. Die Konstruktion der binären Vektoren, die die zu exprimierenden Nukleinsäuren enthalten, erfolgt durch Ligation der
35 cDNA in Sense-Orientierung in die T-DNA erfolgen. 5' der cDNA aktiviert ein Pflanzenpromotor die Transkription der cDNA. Eine Polyadenylierungssequenz befindet sich 3'

von der cDNA. Die binären Vektoren können unterschiedliche Markergene tragen wie beispielsweise das Acetolactat Synthasegens (AHAS oder ALS) [Ott et al., J. Mol. Biol. 1996, 263:359-360], das eine Resistenz gegen die Imidazolinone vermittelt oder das nptII-Markergen, das für eine Kanamycin-Resistenz vermittelt durch Neomycinphosphotransferase codiert.

Die gewebespezifische Expression der Nukleinsäuren lässt sich unter Verwendung eines gewebespezifischen Promotors erzielen. Wenn nicht anders beschrieben wurde der LeB4- oder der USP-Promotor oder der Phaseolin-Promotor 5' der cDNA einkloniert wird. Als Terminatoren wurde der NOS-Terminator und der OCS-Terminator verwendet (siehe Figur 8). Figur 8 zeigt eine Vektorkarte des zur Expression verwendeten Vektor pSUN3CeLPLAT.

Auch jedes andere samenspezifische Promotorelement wie z.B. der Napin- oder Arceolin Promotor Goossens et al. 1999, Plant Phys. 120(4):1095-1103 und Gerhardt et al. 2000, Biochimica et Biophysica Acta 1490(1-2):87-98) kann verwendet werden.

Zur konstitutiven Expression in der ganzen Pflanzen lässt sich der CaMV-35S-Promotor oder ein v-ATPase C1 Promotor verwenden.

Die im Verfahren verwendeten Nukleinsäuren, die für die Acyl-CoA:Lysophospholipid-Acyltransferasen; Desaturasen oder Elongasen codieren, wurden durch Konstruktion mehrerer Expressionskassetten hintereinander in einen binären Vektor kloniert, um den Stoffwechselweg in Pflanzen nachzubilden.

Innerhalb einer Expressionskassette kann das zu exprimierende Protein unter Verwendung eines Signalpeptids, beispielsweise für Plastiden, Mitochondrien oder das Endoplasmatische Retikulum, in ein zelluläres Kompartiment dirigiert werden (Kernode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423). Das Signalpeptid wird 5' im Leseraster mit der cDNA einkloniert, um die subzelluläre Lokalisierung des Fusionsprotein zu erreichen.

Beispiele für Multiexpressionskassetten wurden in DE 102 19 203 offenbart und sind im folgenden nochmals wiedergegeben.

i.) Promotor-Terminator-Kassetten

Expressionskassetten bestehen aus wenigstens zwei funktionellen Einheiten wie einem Promotor und einem Terminator. Zwischen Promotor und Terminator können weitere gewünschte Gensequenzen wie Targetting-Sequenzen, Codierregionen von Genen oder Teilen davon etc. eingefügt werden. Zum Aufbau der Expressionskassetten wurden Promotoren und Terminatoren (USP Promotor: Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67); OCS Terminator: Gielen et al. EMBO J. 3 (1984) 835ff.) mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl auf Basis von synthetischen Oligonukleotiden maßgeschneidert.

Folgende Oligonukleotide können beispielsweise verwendet werden:

USP1 vorne:

- CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA -

USP2 vorne:

5 - CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA -

USP3 vorne:

- CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA -

USP1 hinten:

- AAAACTGCAGGCGGCCGCCACCGCGGTGGGCTGGCTATGAAGAAATT -

10 USP2 hinten:

- CGCGGATCCGCTGGCTATGAAGAAATT -

USP3 hinten:

- TCCCCCGGGATCGATGCCGGCAGATCTGCTGGCTATGAAGAAATT -

OCS1 vorne:

15 - AAAACTGCAGTCTAGAAGGCCTCCTGCTTTAATGAGATAT -

OCS2 vorne:

- CGCGGATCCGATATCGGGCCCGCTAGCGTTAACCCTGCTTTAATGAGATAT -

OCS3 vorne:

- TCCCCCGGGCCATGGCCTGCTTTAATGAGATAT -

20 OCS1 hinten:

- CCCAAGCTTGGCGCGCCGAGCTCGAATTCGTGACGGACAATCAGTAAATTGA -

OCS2 hinten:

- CCCAAGCTTGGCGCGCCGAGCTCGAATTCGTGACGGACAATCAGTAAATTGA -

OCS3 hinten:

25 - CCCAAGCTTGGCGCGCCGAGCTCGTTCGACGGACAATCAGTAAATTGA -

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind allgemein literaturbekannt.

30 In einem ersten Schritt wurden ein Promotor und ein Terminator über PCR amplifiziert. Dann wurde der Terminator in ein Empfängerplasmid kloniert und in einem zweiten Schritt der Promotor vor den Terminator inseriert. Dadurch wurde eine Expressionskassette in das Basis-Plasmid cloniert. Auf Basis des Plasmides pUC19 wurden so die Plasmide pUT1, 2 und 3 erstellt.

Die entsprechenden Konstrukte bzw. Plasmide sind in SEQ ID NO: 15, 16 bis 17 definiert. Sie enthalten den USP-Promotor und den OCS Terminator. Auf Basis dieser Plasmide wurde das Konstrukt pUT12 erstellt, indem pUT1 mittels SalI/Scal geschnitten wurde und pUT2 mittels XhoI/Scal geschnitten wurde. Die die Expressionskassetten enthaltenden Fragmente wurden ligiert und in E. coli XL1 blue MRF transformiert. Es wurde nach Vereinzelung von ampicillinresistenten Kolonien DNA präpariert und per Restriktionsanalyse solche Klone identifiziert, die zwei Expressionskassetten enthalten. Die XhoI/SalI Ligation kompatibler Enden hat dabei die beiden Schnittstellen XhoI und SalI zwischen den Expressionskassetten eliminiert. Das resultierende Plasmid pUT12 wird in SEQ ID NO: 18 wiedergegeben. Anschließend wurde pUT12 wiederum mittels SalI/Scal geschnitten und pUT3 mittels XhoI/Scal geschnitten. Die die Expressionskassetten enthaltenden Fragmente wurden ligiert und in E. coli XLI blue MRF transformiert. Es wurde wieder nach Vereinzelung aus ampicillinresistenten Kolonien DNA präpariert und per Restriktionsanalyse solche Klone identifiziert, die drei Expressionskassetten enthalten. Auf diese Weise wurde ein Set von Multiexpressionskassetten geschaffen, dass für die Insertion gewünschter DNA genutzt werden kann und in Tabelle 1 beschrieben wird und zudem noch weitere Expressionskassetten aufnehmen kann.

Diese enthalten folgende Elemente:

20 Tabelle 1

PUC19-Derivat	Schnittstellen vor dem USP Promotor	Multiple Klonierungs-Schnittstellen	Schnittstellen hinter dem OCS-Terminator
PUT1	EcoRI/AscI/ SacI/XhoI	BstXI/NotI/ PstI/XbaI/StuI	SalI/EcoRI/ SacI/AscI/ HindIII
PUT2	EcoRI/AscI/ SacI/XhoI	BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/ SacI/AscI/ HindIII
PUT3	EcoRI/AscI/ SacI/XhoI	BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/ AscI/HindIII
PUT12 Doppel-expressions-kassette	EcoRI/AscI/ SacI/XhoI	BstXI/NotI/ PstI/XbaI/StuI Und BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/ SacI/AscI/ HindIII
PUT123 Tripel-expressions-kassette	EcoRI/AscI/ SacI/XhoI	1. BstXI/NotI/ PstI/XbaI/StuI und 2. BamHI/EcoRV/ ApaI/NheI/ HpaI und 3. BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/AscI/HindIII

Weiterhin lassen sich wie beschrieben und wie in Tabelle 2 näher spezifiziert weitere Multiexpressionskassetten mithilfe des

- i) USP-Promotors oder mithilfe des
- ii) 700 Basenpaare 3'-Fragmentes des LeB4-Promotors oder mithilfe des
- iii) DC3-Promotors erzeugen und für samenspezifische Genexpression einsetzen.

Der DC3-Promotor ist beschrieben bei Thomas, Plant Cell 1996, 263:359-368 und besteht lediglich aus der Region -117 bis +26 weshalb er mithin einer der kleinsten bekannten samenspezifischen Promotoren darstellt. Die Expressionskassetten können mehrfach den selben Promotor enthalten oder aber über drei verschiedene Promotoren aufgebaut werden.

Vorteilhaft verwendete Polylinker- bzw. Polylinker-Terminator-Polylinker sind den Sequenzen SEQ ID NO: 23 bis 25 zu entnehmen.

Tabelle 2: Multiple Expressionskassetten

Plasmidname des pUC19-Derivates	Schnittstellen vor dem jeweiligen Promotor	Multiple Klonierungs-Schnittstellen	Schnittstellen hinter dem OCS-Terminator
pUT1 (pUC19 mit USP-OCS1)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/PstI/XbaI/StuI	Sall/EcoRI/SacI/AscI/HindIII
PDCT (pUC19 mit DC3-OCS)	EcoRI/AscI/SacI/XhoI	(2) BamHI/EcoRV/ ApaI/NheI/HpaI	Sall/EcoRI/SacI/AscI/HindIII
PleBT (pUC19-mit LeB4(700)-OCS)	EcoRI/AscI/SacI/XhoI	(3) BglII/NaeI/ ClaI/SmaI/NcoI	Sall/SacI/AscI/HindIII
PUD12 (pUC 19 mit mit USP-OCS1 und mit DC3-OCS)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/ PstI/XbaI/StuI und (2) BamHI/EcoRV/ ApaI/NheI/HpaI	Sall/EcoRI/SacI/AscI/HindIII
PUDL123 Triple expression cassette (pUC19 mit USP/DC3 und LeB4-700)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/ PstI/XbaI/StuI und (2) BamHI/ (EcoRV*)/ApaI/NheI/HpaI und (3) BglII/NaeI/ ClaI/SmaI/NcoI	Sall/SacI/AscI/HindIII

15. * EcoRV Schnittstelle schneidet im 700 Basenpaarfragment des LeB4 Promotors (LeB4-700)

Analog lassen sich weitere Promotoren für Multigenkonstrukte erzeugen insbesondere unter Verwendung des

- a) 2,7 kB Fragmentes des LeB4-Promotors oder mithilfe des
- b) Phaseolin-Promotors oder mithilfe des
- 5 c) konstitutiven v-ATPase c1-Promotors.

Es kann insbesondere wünschenswert sein, weitere besonders geeignete Promotoren zum Aufbau samenspezifischer Multiexpressionskassetten wie z.B. den Napin-Promotor oder den Arcelin-5 Promotor zu verwenden.

- 10 Weitere in Pflanzen nutzbare Vektoren mit einer bzw. zwei oder drei Promotor-Terminator-Expressionkassetten sind den Sequenzen SEQ ID NO: 26 bis SEQ ID NO: 31 zu entnehmen.

- ii.) Erstellung von Expressionskonstrukten, die Promotor, Terminator und gewünschte Gensequenz zur PUFA Genexpression in pflanzlichen Expressionskassetten enthalten.

- 15 In pUT123 wird zunächst über BstXI und XbaI die Δ -6-Elongase Pp_PSE1 in die erste Kasette inseriert. Dann wird die Δ -6-Desaturase aus Moos (Pp_des6) über BamHI/NaeI in die zweite Kasette inseriert und schließlich die Δ -5-Desaturase aus Phaeodactylum (Pt_des5) über BglII/NcoI in die dritte Kasette inseriert (siehe SEQ ID NO: 19). Das Dreifachkonstrukt erhält den Namen pARA1. Unter Berücksichtigung sequenzspezifischer Restriktionsschnittstellen können weitere Expressionskassetten gemäß Tabelle 3 mit der Bezeichnung pARA2, pARA3 und pARA4 erstellt werden.
- 20

Tabelle 3: Kombinationen von Desaturasen und Elongasen

Gen Plasmid	Δ -6-Desaturase	Δ -5-Desaturase	Δ -6-Elongase
pARA1	Pp_des6	Pt_des5	Pp_PSE1
pARA2	Pt_des6	Pt_des5	Pp_PSE1
pARA3	Pt_des6	Ce_des5	Pp_PSE1
PARA4	Ce_des6	Ce_des5	Ce_PSE1

des5 = PUFA spezifische Δ -5-Desaturase

des6 = PUFA spezifische Δ -6-Desaturase

5 PSE = PUFA spezifische Δ -6-Elongase

Pt_des5 = Δ -5-Desaturase aus *Phaeodactylum tricornutum*

Pp_des6 oder Pt_des6 = Δ -6-Desaturase aus *Physcomitrella patens* bzw. *Phaeodactylum tricornutum*

Pp = *Physcomitrella patens*, Pt = *Phaeodactylum tricornutum*

10 Pp_PSE1 = Δ -6-Elongase aus *Physcomitrella patens*

Pt_PSE1 = Δ -6-Elongase aus *Phaeodactylum tricornutum*

Ce_des5 = Δ -5-Desaturase aus *Caenorhabditis elegans* (Genbank Acc. Nr. AF078796)

Ce_des6 = Δ -6-Desaturase aus *Caenorhabditis elegans* (Genbank Acc. Nr. AF031477, Basen 11-1342)

15 Ce_PSE1 = Δ -6-Elongase aus *Caenorhabditis elegans* (Genbank Acc. Nr. AF244356, Basen 1-867)

Auch weitere Desaturasen oder Elongasegensequenzen können in Expressionskassetten beschriebener Art inseriert werden wie z.B. Genbank Acc. Nr. AF231981, NM_013402, AF206662, AF268031, AF226273, AF110510 oder AF110509.

20 iii.) Transfer von Expressionskassetten in Vektoren zur Transformation von *Agrobacterium tumefaciens* und zur Transformation von Pflanzen

Die so erstellten Konstrukte wurden mittels *AscI* in den binären Vektor pGPTV inseriert. Die multiple Klonierungssequenz wurde zu diesem Zweck um eine *AscI* Schnittstelle erweitert. Zu diesem Zweck wurde der Polylinker als zwei doppelsträngige Oligonukleotide neu synthetisiert, wobei eine zusätzliche *AscI* DNA Sequenz eingefügt wird. Das Oligonukleotid wurde mittels *EcoRI* und *HindIII* in den Vektor pGPTV inseriert. Die notwendigen Kloniertechniken sind dem Fachmann bekannt und können einfach wie in Beispiel 1 beschrieben nachgelesen werden.

30 Für die im folgenden beschriebenen Versuche wurden als Nukleinsäuresequenzen für die Δ -5-Desaturase (SEQ ID NO: 13), die Δ -6-Desaturase (SEQ ID NO: 9) und die Δ -6-Elongase (SEQ ID NO: 11), die Sequenzen aus *Physcomitrella patens* und *Phaeodactylum tricornutum* verwendet. Die entsprechenden Aminosäuresequenzen sind den Sequenzen SEQ ID NO: 10, SEQ ID NO: 12 und SEQ ID NO: 14. Ein Vektor der alle vorgenannten Gene enthält ist in SEQ ID NO: 19 wiedergegeben. Die korrespondierenden

Aminosäuresequenzen der Gene sind SEQ ID NO: 20, SEQ ID NO: 21 und SEQ ID NO: 22 zu entnehmen.

Beispiel 2: Klonierung und Charakterisierung der ceLPLATs

a) Datenbanken-Suche

- 5 Die Identifizierung der ceLPLATs (= Acyl-CoA:Lysophospholipid-Acyltransferase aus *Caenorhabditis elegans*) erfolgte durch Sequenzvergleiche mit bekannten LPA-ATs. Die Suche wurde mit Hilfe des BLAST-Psi-Algorithmus (Altschul et al., J. Mol. Biol. 1990, 215: 403-410) auf das Nematodengenom (*Caenorhabditis elegans*) beschränkt, da dieser Organismus LCPUFAs synthetisiert. Für den Sequenzvergleich diente als
- 10 Sonde eine LPAAT Proteinsequenz aus *Mus musculus* (MsLPAAT Accession Nr. NP_061350). LPLAT katalysiert durch eine reversible Transferasereaktion die ATP-unabhängige Synthese von Acyl-CoAs aus Phospholipiden mit Hilfe von CoA als Co-factor (Yamashita et al., J. Biol. Chem. 2001, 20: 26745-26752). Durch Sequenzvergleiche konnten zwei putative ceLPLAT-Sequenzen identifiziert werden (Accession Nr.
- 15 T06E8.1 bzw. F59F4.4). Die identifizierten Sequenzen weisen die größte Ähnlichkeit jeweils zueinander und zu MsLPAATs auf (Figur 1). Das Alignment wurde mit dem Programm Clustal erstellt.

b) Klonierung der CeLPLATs

- 20 Auf der Basis der ceLPLAT-Nukleinsäuresequenzen wurden Primerpaare synthetisiert (Tab. 1) und mittels PCR-Verfahren die zugehörigen cDNAs aus einer *C. elegans*-cDNA-Bank isoliert. Die entsprechenden Primerpaare wurden so ausgewählt, dass sie die Hefe-Konsensus-Sequenz für hocheffiziente Translation (Kozak, Cell 1986, 44:283-292) neben dem Startcodon trugen. Die Amplifizierung der LPLAT-cDNAs wurde je-
- 25 weils mit 2 µl cDNA-Bank-Lösung als Template, 200 µM dNTPs, 2,5 U "proof-reading" *pfu*-Polymerase und 50 pmol eines jeden Primers in einem Gesamtvolumen von 50 µl durchgeführt. Die Bedingungen für die PCR waren wie folgt: Erste Denaturierung bei 95°C für 5 Minuten, gefolgt von 30 Zyklen bei 94°C für 30 Sekunden, 58°C für eine Minute und 72°C für 2 Minuten sowie ein letzter Verlängerungsschritt bei 72°C für 10 Minuten. Die Sequenz der LPLAT-cDNAs wurde durch DNA-Sequenzierung bestätigt.

Tabelle 4: Nukleotidsequenzen der PCR-Primer zur Klonierung von CeLPLATs

Primer	Nukleotidsequenz
5' T06E8.1f*	5' ACATAATGGAGAACTTCTGGTCGATCGTC 3'
3' T06E8.1r*	5' TTA CTCAGATTTCTTCCCGTCTTT 3'
5' F59F4.4f*	5' ACATAATGACCTTCCTAGCCATATTA 3'
3' F59F4.4r*	5' TCAGATATTCAAATTGGCGGCTTC 3'

* f: forward, r: reverse

Beispiel 3: Analyse der Auswirkung der rekombinanten Proteine auf die Produktion des gewünschten Produktes

5 a) Aufarbeitungsmöglichkeiten

Die Auswirkung der genetischen Modifikation in Pilzen, Algen, Ciliaten oder wie in den Beispielen weiter oben beschrieben in Hefen auf die Produktion der mehrfach ungesättigten Fettsäuren oder Pflanzen kann bestimmt werden, indem die modifizierten Mikroorganismen oder die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen) gezüchtet werden und das Medium und/oder die zellulären Komponenten auf die erhöhte Produktion der Lipide oder Fettsäuren untersucht wird. Diese Analysetechniken sind dem Fachmann bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie (siehe beispielsweise Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et al. (1993) Biotechnology, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988) Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

Neben den oben erwähnten Verfahren zum Nachweis von Fettsäuren in Hefen werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-

145, beschrieben extrahiert. Die qualitative und quantitative Lipid- oder Fettsäureanalyse ist beschrieben bei Christie, William W., *Advances in Lipid Methodology*, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., *Gas Chromatography and Lipids. A Practical Guide* - Ayr, Scotland: Oily Press, 1989, Repr. 1992, 1X, 307 S. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

10 So kann die Analyse von Fettsäuren oder Triacylglycerin (= TAG, Abkürzungen in Klammern angegeben) z.B. mittels Fettsäuremethylester (= FAME), Gas-Flüssigkeitschromatographie-Massenspektrometrie (= GC-MS) oder Dünnschichtchromatographie (TLC) erfolgen.

15 Der unzweideutige Nachweis für das Vorliegen von Fettsäureprodukten kann mittels Analyse rekombinanter Organismen nach Standard-Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: *Advances on Lipid Methodology*, Vierte Aufl.: Christie, Oily Press, Dundee, 119-169; 1998, *Gaschromatographie-Massenspektrometrie-Verfahren*, *Lipide* 33:343-353).

20 Das zu analysierende Pflanzenmaterial kann dazu entweder durch Ultraschallbehandlung, Mahlen in der Glasmühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren aufgebrochen werden. Das Material wird dann anschließend nach dem Aufbrechen zentrifugiert. Das Sediment wird danach in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit 2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl- und Lipidverbindungen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester können anschließend in Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapillarsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 Mikrom, 0,32 mm) bei einem Temperaturgradienten zwischen 170°C und 240°C für 20 min und 5 min bei 240°C unterworfen werden. Die Identität der erhaltenen Fettsäuremethylester lassen sich unter Verwendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma), definieren.

35 Bei Fettsäuren, für die keine Standards verfügbar sind, kann die Identität über Derivatisierung und anschließende GC-MS-Analyse gezeigt werden. Beispielsweise wird die Lokalisierung von Fettsäuren mit Dreifachbindung über GC-MS nach Derivatisierung mit 4,4-Dimethoxyoxazolin-Derivaten (Christie, 1998, siehe oben) gezeigt.

b) Fettsäureanalyse in Pflanzen

Die Gesamt-Fettsäuren wurden aus Pflanzensamen extrahiert und mittels Gaschromatographie analysiert.

Die Samen wurden mit 1 % Natriummethanolat in Methoanol aufgenommen und 20 min bei RT (ca. 22 °C) inkubiert. Anschließend wurde mit NaCl Lösung gewaschen und die FAME in 0,3 ml Heptan aufgenommen.

- Die Proben wurden auf einer ZEBRON-ZB-Wax-Kapillarsäule (30 m, 0,32 mm, 0,25 mikro m; Phenomenex) in einem Hewlett Packard-6850-Gaschromatograph mit einem Flammenionisationsdetektor aufgetrennt. Die Ofentemperatur wurde von 70°C (1 min halten) bis 200°C mit einer Rate von 20°C/min, dann auf 250°C (5 min halten) mit einer Rate von 5°C/min und schließlich auf 260°C mit einer Rate von 5°C/min programmiert. Stickstoff wurde als Trägergas verwendet (4,5 ml/min bei 70°C). Die Fettsäuren wurden durch Vergleich mit Retentionszeiten von FAME-Standards (SIGMA) identifiziert.

Beispiel 4: Funktionelle Charakterisierung der CeLPLATs in Hefe

a) Heterologe Expression in *Saccharomyces cerevisiae*

- Zur Charakterisierung der Funktion der CeLPLATs aus *C. elegans* wurden die offenen Leserahmen der jeweiligen cDNAs stromabwärts des Galactose-induzierbaren GAL1-Promotors von pYes2.1Topo unter Verwendung des pYes2.1TOPO TA Expression Kit (Invitrogen) kloniert, wobei pYes2-T06E8.1 und pYes2-F59F4.4 erhalten wurden.

- Da die Expression der CeLPLATs zu einem effizienten Austausch der Acyl-Substrate führen sollte, wurde weiterhin das Doppelkonstrukt pESCLeu-PpD6-Pse1 hergestellt, das die offenen Leserahmen einer $\Delta 6$ -Desaturase (PpD6) und einer $\Delta 6$ -Elongase (PSE1) aus *Physcomitrella patens* (siehe DE 102 19 203) beinhaltet. Die Nukleinsäuresequenz der $\Delta 6$ -Desaturase (PpD6) und der $\Delta 6$ -Elongase (Pse1) werden jeweils in SEQ ID NO: 9 und SEQ ID NO: 11 wiedergegeben. Die korrespondierenden Aminosäuresequenzen sind SEQ ID NO: 10 und SEQ ID NO: 12 zu entnehmen.

- Die *Saccharomyces cerevisiae*-Stämme C13ABYS86 (Protease-defizient) und INVSc1 wurde mittels eines modifizierten PEG/Lithiumacetat-Protokolls gleichzeitig mit den Vektoren pYes2-T06E8.1 und pESCLeu-PpD6-Pse1 bzw. pYes2-F59F4.4 und pESCLeu-PpD6-Pse1 transformiert. Als Kontrolle wurde eine Hefe verwendet, die mit dem Vektor pESCLeu-PpD6-Pse1 und dem leeren Vektor pYes2 transformiert wurde. Die Selektion der transformierten Hefen erfolgte auf Komplett-Minimalmedium (CMdum)-Agarplatten mit 2% Glucose, aber ohne Uracil und Leucin. Nach der Selektion wurden 4 Transformanten, zwei pYes2-T06E8.1/pESCLeu-PpD6-Pse1 und zwei pYes2-F59F4.4/pESCLeu-PpD6-Pse1 und eine pESCLeu-PpD6-Pse1/ pYes2 zur weiteren funktionellen Expression ausgewählt. Die beschriebenen Experimente wurden auch im Hefestamm INVSc1 durchgeführt.

Für die Expression der CeLPAATs wurden zunächst Vorkulturen aus jeweils 2 ml CMdum-Flüssigmedium mit 2% (w/v) Raffinose, aber ohne Uracil und Leucin mit den ausgewählten Transformanten angeimpft und 2 Tage bei 30°C, 200rpm inkubiert. 5 ml CMdum-Flüssigmedium (ohne Uracil und Leucin) mit 2% Raffinose, 1% (v/v) Tergitol

NP-40 und 250 μ M Linolsäure ($18:2^{\Delta 9,12}$) oder Linolensäure ($18:3^{\Delta 9,12,15}$) wurden dann mit den Vorkulturen auf eine OD_{600} von 0,08 angeimpft. Die Expression wurde bei einer OD_{600} von 0,2-0,4 durch die Zugabe von 2% (w/v) Galactose induziert. Die Kulturen wurden für weitere 48 h bei 20°C inkubiert.

5 Fettsäureanalyse

- Die Hefezellen aus den Hauptkulturen wurden durch Zentrifugation (100 x g, 10 min, 20°C) geerntet und mit 100 mM NaHCO_3 , pH 8,0 gewaschen, um restliches Medium und Fettsäuren zu entfernen. Aus den Hefe-Zellsedimenten wurden Fettsäuremethylester (FAMES) durch saure Methanolyse hergestellt. Hierzu wurden die Zellsedimente mit 2 ml 1N methanolischer Schwefelsäure und 2% (v/v) Dimethoxypropan für 1 h bei 80°C inkubiert. Die Extraktion der FAMES erfolgte durch zweimalige Extraktion mit Petrolether (PE). Zur Entfernung nicht derivatisierter Fettsäuren wurden die organischen Phasen je einmal mit 2 ml 100 mM NaHCO_3 , pH 8,0 und 2 ml Aqua dest. gewaschen. Anschließend wurden die PE-Phasen mit Na_2SO_4 getrocknet, unter Argon eingedampft und in 100 μ l PE aufgenommen. Die Proben wurden auf einer DB-23-Kapillarsäule (30 m, 0,25 mm, 0,25 μ m, Agilent) in einem Hewlett-Packard 6850-Gaschromatographen mit Flammenionisationsdetektor getrennt. Die Bedingungen für die GLC-Analyse waren wie folgt: Die Ofentemperatur wurde von 50°C bis 250°C mit einer Rate 5°C/min und schließlich 10 min bei 250°C (halten) programmiert.
- Die Identifikation der Signale erfolgte durch Vergleiche der Retentionszeiten mit entsprechenden Fettsäurestandards (Sigma).

Acyl-CoA Analyse

Die Acyl-CoA-Analyse erfolgte wie bei Larson and Graham (2001; Plant Journal 25: 115-125) beschrieben.

25 Expressionsanalyse

- Figuren 2 A und B sowie 3 A und B zeigen die Fettsäureprofile von transgenen C13ABYS86 Hefen, die mit $18:2^{\Delta 9,12}$ bzw. $18:3^{\Delta 9,12,15}$ gefüttert wurden. Die gefütterten Substrate sind in großen Mengen in allen transgenen Hefen nachzuweisen. Alle vier transgenen Hefen zeigen eine Synthese von $18:3^{\Delta 6,9,12}$ und $20:3^{\Delta 8,11,14}$ bzw. $18:4^{\Delta 6,9,12,15}$ und $20:4^{\Delta 8,11,14,17}$, den Produkten der Δ -6-Desaturase und Δ -6-Elongase Reaktionen. Dies bedeutet, dass die Gene PpD6 und Pse1 funktional exprimiert werden konnten.

- Figur 2 gibt wie oben beschrieben die Fettsäureprofile von transgenen C13ABYS86 *S. cerevisiae*-Zellen. Die Synthese der Fettsäuremethylester erfolgte durch saure Methanolyse intakter Zellen, die entweder mit den Vektoren pESCLeu-PpD6-Pse1/pYes2 (A) oder pYes2-T06E8.1/pESCLeu-PpD6-Pse1 (B) transformiert worden waren. Die Hefen wurden in Minimalmedium in Gegenwart von $18:2^{\Delta 9,12}$ kultiviert. Anschließend wurden die Fettsäuremethylester über GLC analysiert.

- In den Kontroll-Hefen, die mit den Vektoren pESCLeu-PpD6-Pse1/pYes2 transformiert wurden, ist der Anteil von 20:3^{Δ8,11,14}, zu dem 18:3^{Δ6,9,12} durch Pse1 elongiert wird, wesentlich niedriger als in den Hefen, die zusätzlich die LPLAT T06E8.1 exprimieren. Tatsächlich konnte die Elongation von 18:3^{Δ6,9,12} und 18:4^{Δ6,9,12,15} durch die zusätzliche Expression von CeLPLAT (T06E8.1) um 100-150% verbessert werden (Figur 4). Diese signifikante Erhöhung des LCPUFA-Gehalts ist nur wie folgt zu erklären: die exogen gefütterten Fettsäuren (18:2^{Δ9,12} bzw. 18:3^{Δ9,12,15}) werden zunächst in Phospholipide eingebaut und dort von der Δ-6-Desaturase zu 18:3^{Δ6,9,12} und 18:4^{Δ6,9,12,15} desaturiert. Erst nach Reäquilibration mit dem Acyl-CoA-Pool können 18:3^{Δ6,9,12} und 18:4^{Δ6,9,12,15} durch die Elongase zu 20:3^{Δ8,11,14} bzw. 20:4^{Δ8,11,14,17}-CoA elongiert und dann wieder in die Lipide eingebaut werden. Die LPLAT T06E8.1 ist in der Lage, die Δ6-desaturierten Acylgruppen sehr effizient in CoA-Thioester zurückzuverwandeln. Interessanterweise konnte auch die Elongation der gefütterten Fettsäuren 18:2^{Δ9,12} und 18:3^{Δ9,12,15} verbessert werden. (Figur 2 A und B bzw. 3 A und B).
- Figur 3 gibt die Fettsäureprofile von transgenen C13ABYS86 *S. cerevisiae*-Zellen. Die Synthese der Fettsäuremethylester erfolgte durch saure Methanolyse intakter Zellen, die entweder mit den Vektoren pESCLeu-PpD6-Pse1/pYes2 (A) oder pYes2-T06E8.1/pESCLeu-PpD6-Pse1 (B) transformiert worden waren. Die Hefen wurden in Minimalmedium in Gegenwart von 18:3^{Δ9,12,15} kultiviert. Anschließend wurden die Fettsäuremethylester über GLC analysiert.

Die Expression einer anderen CeLPLAT (F59F4.4) hat dagegen keinen Einfluss auf die Elongation (Figur 4). Offenbar kodiert F59F4.4 nicht für eine LPLAT. Nicht jede der putativen LPLAT Nukleinsäuresequenzen ist also enzymatisch aktiv in der erfindungsgemäß gefundenen Reaktion.

- Figur 4 gibt die Elongation exogen applizierter 18:2^{Δ9,12} bzw. 18:3^{Δ9,12,15} im Anschluss an ihre endogene Δ-6-Desaturierung (Daten aus Fig. 2 und 3) wieder. Die exogen gefütterten Fettsäuren werden zunächst in Phospholipide eingebaut und dort zu 18:3^{Δ6,9,12} und 18:4^{Δ6,9,12,15} desaturiert. Erst nach Reäquilibration mit dem Acyl-CoA-Pool können 18:3^{Δ6,9,12} und 18:4^{Δ6,9,12,15} durch die Elongase zu 20:3^{Δ8,11,14} bzw. 20:4^{Δ8,11,14,17}-CoA elongiert und dann wieder in die Lipide eingebaut werden. Die LPLAT T06E8.1 ist in der Lage, die Δ-6-desaturierten Acylgruppen effizient in CoA-Thioester zurückzuverwandeln.

- Diese Ergebnisse zeigen, dass die CeLPLAT (T06E8.1) nach Co-expression mit der Δ-6-Desaturase und Δ-6-Elongase zu einer effizienten Produktion von C20-PUFAs führt. Diese Ergebnisse sind dadurch zu erklären, dass die CeLPLAT (T06E8.1) einen effizienten Austausch der neusynthetisierten Fettsäuren zwischen Lipiden und dem Acyl-CoA-Pool ermöglicht (siehe Figur 7).

- Figur 7 gibt die Acyl-CoA-Zusammensetzung transgener INVSc1 Hefen, die mit den Vektoren pESCLeu PpD6Pse1/pYes2 (A) oder pESCLeu-PpD6-Pse1/pYes2-T06E8.1 (B) transformiert worden waren, wieder. Die Hefezellen wurden in Minimalmedium oh-

ne Uracil und Leucin in Gegenwart von 250 μ M 18:2 ^{Δ 9,12} kultiviert. Die Acyl-CoA-Derivate wurden über HPLC analysiert.

Bei Verwendung des Hefe-Stammes INVSc1 zur Co-Expression von CeLPLAT (T06E8.1) zusammen mit PpD6 und Pse1 ergibt sich folgendes Bild: Kontrollhefen, die PpD6 und Pse1 exprimieren, enthalten wie schon bei Verwendung des Stammes C13ABYS86 gezeigt nur geringe Mengen des Elongationsprodukts (20:3 ^{Δ 8,11,14} bei Fütterung von 18:2 bzw. 20:4 ^{Δ 8,11,14,17} bei Fütterung von 18:3; siehe Figur 5 A und 6 A). Bei zusätzlicher Expression von CeLPLAT (T06E8.1) erfolgt ein deutlicher Anstieg dieser Elongationsprodukte (siehe Figur 5 B und 6 B). Tabelle 6 zeigt, dass die zusätzliche Expression von CeLPLAT überraschenderweise eine 8-fache Erhöhung des Gehaltes an 20:3 ^{Δ 8,11,14} (bei Fütterung von 18:2) bzw. 20:4 ^{Δ 8,11,14,17} (bei Fütterung von 18:3) bewirkt. Daneben zeigt sich, dass auch C16:2 ^{Δ 6,9} zu C18:2 ^{Δ 6,9} effizienter elongiert wird.

Tabelle 5: Fettsäure-Zusammensetzung (in mol %) transgener Hefen, die mit den Vektoren pESCLeu PpD6Pse1/pYes2 (PpD6 Pse1) oder pESCLeu-PpD6-Pse1/pYes2-T06E8.1 (PpD6 Pse1 + T06E8) transformiert worden waren. Die Hefezellen wurden in Minimalmedium ohne Uracil und Leucin in Gegenwart von 250 μ M 18:2 $\Delta^{9,12}$ oder 18:3 $\Delta^{9,12,15}$ kultiviert. Die Fettsäuremethylester wurden durch saure Methanolyse ganzer Zellen gewonnen und über GLC analysiert. Jeder Wert gibt den Mittelwert (n = 4) \pm Standardabweichung wieder.

Fettsäuren	Fütterung mit 250 μ M 18:2 $\Delta^{9,12}$		Fütterung mit 250 μ M 18:3 $\Delta^{9,12,15}$	
	PpD6/Pse1	PpD6/Pse1+	PpD6/Pse1	PpD6/Pse1+
		T06E8		T06E8
16:0	15,31 \pm 1,36	15,60 \pm 1,36	12,20 \pm 0,62	16,25 \pm 1,85
16:1 Δ^9	23,22 \pm 2,16	15,80 \pm 3,92	17,61 \pm 1,05	14,58 \pm 1,93
18:0	5,11 \pm 0,63	7,98 \pm 1,28	5,94 \pm 0,71	7,52 \pm 0,89
18:1 Δ^9	15,09 \pm 0,59	16,01 \pm 2,53	15,62 \pm 0,34	15,14 \pm 2,61
18:1 Δ^{11}	4,64 \pm 1,09	11,80 \pm 1,12	4,56 \pm 0,18	13,07 \pm 1,66
18:2 $\Delta^{9,12}$	28,72 \pm 3,25	14,44 \pm 1,61	-	-
18:3 $\Delta^{6,9,12}$	3,77 \pm 0,41	4,72 \pm 0,72	-	-
18:3 $\Delta^{9,12,15}$	-	-	32,86 \pm 1,20	14,14 \pm 2,52
18:4 $\Delta^{6,9,12,15}$	-	-	5,16 \pm 1,04	3,31 \pm 1,15
20:2 $\Delta^{11,14}$	2,12 \pm 0,86	4,95 \pm 4,71	-	-
20:3 $\Delta^{8,11,14}$	1,03 \pm 0,14	8,23 \pm 1,59	-	-
20:3 $\Delta^{11,14,17}$	-	-	4,12 \pm 1,54	6,95 \pm 2,52
20:4 $\Delta^{8,11,14,17}$	-	-	1,34 \pm 0,28	8,70 \pm 1,11

10~ Fig. 5 ist das Fettsäure-Profil von transgenen INVSc1 *S. cerevisiae*-Zellen zu entnehmen. Die Synthese der Fettsäuremethylester erfolgte durch saure Methanolyse intakter Zellen, die entweder mit den Vektoren pESCLeu-PpD6-Pse1/pYes2 (A) oder pYes2-T06E8.1/pESCLeu-PpD6-Pse1 (B) transformiert worden waren. Die Hefen wur-

den in Minimalmedium in Gegenwart von $18:2^{\Delta 9,12}$ kultiviert. Anschließend wurden die Fettsäuremethylester über GLC analysiert.

Figur 6 gibt die Fettsäure-Profile von transgenen INVSc1 *S. cerevisiae*-Zellen wieder. Die Synthese der Fettsäuremethylester erfolgte durch saure Methanolyse intakter Zellen, die entweder mit den Vektoren pESCLeu-PpD6-Pse1/pYes2 (A) oder pYes2-T06E8.1/pESCLeu-PpD6-Pse1 (B) transformiert worden waren. Die Hefen wurden in Minimalmedium in Gegenwart von $18:3^{\Delta 12,15}$ kultiviert. Anschließend wurden die Fettsäuremethylester über GLC analysiert.

Ein Maß für die Effizienz der LCPUFA-Biosynthese in transgener Hefe stellt der Quotient aus Gehalt der erwünschten Δ -6-Elongationsprodukt nach Δ -6-Desaturierung ($20:3^{\Delta 8,11,14}$ bzw. $20:4^{\Delta 8,11,14,17}$) zu Gehalt an zugeführter Fettsäure ($18:2^{\Delta 9,12}$ bzw. $18:3^{\Delta 9,12,15}$) dar. Dieser Quotient beträgt 0,04 in INVSc1 Kontrollhefen, die PpD6 und Pse1 exprimieren, und 0,60 in Hefen die zusätzlich zu PpD6 und Pse1 CeLPLAT exprimieren. In anderen Worten: der Gehalt an erwünschtem Δ -6-Elongationsprodukt nach Δ -6-Desaturierung bei Co-Expression von CeLPLAT beträgt 60% des Gehalts der jeweils zugeführten Fettsäure. In Kontrollhefen beträgt dieser Gehalt nur ca. 4%. Dies bedeutet eine 15-fache Erhöhung der Effizienz der LCPUFA-Biosynthese in transgener Hefe durch Co-Expression von LPLAT.

Interessanterweise bewirkt die Co-Expression von CeLPLAT nicht nur eine Erhöhung der genannten Elongationsprodukte $20:3^{\Delta 8,11,14}$ bzw. $20:4^{\Delta 8,11,14,17}$, sondern auch eine Erhöhung des Verhältnisses $20:3^{\Delta 8,11,14} : 20:2^{\Delta 11,14}$ bzw. $20:4^{\Delta 8,11,14,17} : 20:3^{\Delta 11,14,17}$. Dies bedeutet, dass in Anwesenheit der LPLAT die Δ -6-Elongase bevorzugt mehrfach ungesättigte Fettsäuren ($18:3^{\Delta 6,9,12}$ und $18:4^{\Delta 6,9,12,15}$) als Substrat verwendet, während bei Abwesenheit der LPLAT keine ausgeprägte Substratspezifität zu erkennen ist (auch $18:2^{\Delta 9,12}$ und $18:3^{\Delta 9,12,15}$ werden elongiert). Grund hierfür können Protein-Protein-Wechselwirkungen zwischen Δ -6-Elongase, Δ -6-Desaturase und LPLAT oder posttranslationale Modifikationen (z.B. partielle Proteolyse) sein. Dies würde auch erklären, warum der oben beschriebene Anstieg von Δ -6-Elongationsprodukten bei Co-Expression von Δ -6-Desaturase, Δ -6-Elongase und LPLAT bei Verwendung eines protease-defizienten Hefestamms geringer ausfällt.

Acyl-CoA Analysen von transgenen INVSc1 Hefen, die mit $18:2^{\Delta 9,12}$ gefüttert wurden, ergaben folgendes Ergebnis: in Kontrollhefen, die PpD6 und Pse1 exprimieren, ist kein $18:3^{\Delta 6,9,12}$ -CoA und $20:3^{\Delta 8,11,14}$ -CoA nachweisbar. Dies weist darauf hin, dass weder das Substrat ($18:3^{\Delta 6,9,12}$ -CoA) noch das Produkt ($20:3^{\Delta 8,11,14}$ -CoA) der Δ -6-Elongase in Kontrollhefen in nachweisbaren Mengen vorhanden ist. Dies lässt darauf schließen, dass der Transfer von $18:3^{\Delta 6,9,12}$ aus Membranlipiden in den Acyl-CoA Pool nicht oder nicht richtig stattfindet. Das bedeutet, dass kaum Substrat für die vorhandene Δ -6-Elongase zur Verfügung steht, was wiederum den geringen Gehalt an Elongationsprodukt in Kontrollhefen erklärt. INVSc1 Hefen, die zusätzlich zur PpD6 und Pse1 die CeLPLAT exprimieren und mit $18:2^{\Delta 9,12}$ gefüttert worden waren, weisen keine signifikanten Mengen an $18:3^{\Delta 6,9,12}$ -CoA auf, wohl aber $20:3^{\Delta 8,11,14}$ -CoA. Dies deutet darauf

hin, dass LPLAT sehr effizient $18:3^{\Delta 6,9,12}$ aus den Membranlipiden in den Acyl-CoA-Pool überführt. $18:3^{\Delta 6,9,12}$ -CoA wird dann von der Δ -6-Elongase elongiert, so dass kein $18:3^{\Delta 6,9,12}$ -CoA, wohl aber $20:3^{\Delta 8,11,14}$ -CoA nachweisbar ist.

b) Funktionelle Charakterisierung der CeLPLATs in transgenen Pflanzen

5. Expression funktionaler CeLPLAT in transgenen Pflanzen

In DE 102 19 203 wurden transgene Pflanzen beschrieben, deren Samenöl durch samenspezifische Expression funktioneller Gene kodierend für Δ -6-Desaturase, Δ -6-Elongase und Δ -5-Desaturase geringe Mengen an ARA und EPA enthält. Der zur Transformation dieser Pflanzen benutzte Vektor ist SEQ ID NO: 19 zu entnehmen. Um den Gehalt an diesen LCPUFAs zu erhöhen, wurde in den genannten transgenen Pflanzen zusätzlich das Gen CeLPLAT (T06E8.1) in Samen exprimiert.

Zu diesem Zweck wurde der kodierende Bereich von CeLPLAT über PCR amplifiziert.

In Tabelle 6 sind die Primer wiedergegeben, die zur Klonierung eines weiteren Clones der ceLPLAT in binäre Vektoren verwendet wurden.

15 Tabelle 6: Nukleotidsequenzen der PCR-Primer zur Klonierung von CeLPLAT (T06E8.1) in den binären Vektor pSUN3

Primer	Nukleotidsequenz
ARe503f*	5' TTAAGCGCGGCCGCATGGAGAACTTCTGGTTCG 3'
ARe504r*	5' ACCTCGGCGGCCGCCCTTTTACTCAGATTTC 3'

* f: forward, r: reverse

Das PCR-Produkt wurde in einen pENTRY Vektor zwischen USP Promotor und OCS-Terminator kloniert. Anschließend wurde die Expressionskassette in die binären Vektoren pSUN300 kloniert. Der entstandene Vektor wurde mit pSUN3CeLPLAT (Figur 8) bezeichnet. Darüber hinaus wurde der kodierende Bereiche von CeLPLAT amplifiziert und zwischen LegB4 Promotor und OCS-Terminator kloniert. Dieser Vektor wurde mit pGPTVCeLPLAT bezeichnet (Figur 9A).

25 Darüberhinaus wurde der kodierende Bereich von CeLPLAT über PCR amplifiziert und zwischen LegB4 Promotor und OCS-Terminator kloniert. Die hierfür verwendeten PCR Primer wurden so ausgewählt, dass in das PCR-Produkt eine effiziente Kosaksequenz eingeführt wurde. Außerdem wurde die DNA-Sequenz von CeLPLAT so verändert, dass sie der codon usage von höheren Pflanzen angepasst war.

Folgende Primer wurden für die PCR verwendet:

Forward primer: 5'-ACATAATGGAGAACTTCTGGTCTATTGTTGTGTTTTTCTA-3'

Reverse primer: 5'-CTAGCTAGCTTACTCAGATTTCTTCCCGTCTTTTGTTC-3'

Das PCR Produkt wurde in den Klonierungsvektor pCR Script kloniert und über die Restriktionsenzyme XmaI und SacI in den Vektor pGPTV LegB4-700 kloniert. Das entstandene Plasmid wurde mit pGPTV LegB4-700 + T06E8.1 bezeichnet (Figur 9A).

Das gleiche PCR Produkt wurde darüber hinaus in einen Multigen-Expressionsvektor kloniert, der bereits die Gene für eine Delta-6-Desaturase aus *Phaeodactylum tricornutum* (SEQ ID NO: 32, Aminosäuresequenz SEQ ID NO: 33) und einer Delta-6-Elongase aus *P. patens* enthielt. Das entstandene Plasmid wurde mit pGPTV USP/OCS-1,2,3 PSE1(Pp)+D6-Des(Pt)+2AT (T06E8-1) bezeichnet (Figur 9B). Die Sequenzen des Vektors sowie der Gene sind SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36 und SEQ ID NO: 37 zu entnehmen. Die Δ -6-Desaturase aus *Phaeodactylum tricornutum* reicht von Nukleotid 4554 bis 5987 in der SEQ ID NO: 34. Die Δ -6-Elongase aus *Physcomitrella patens* reicht von Nukleotid 1026 bis 1898 und die der LPLAT aus *Caenorhabditis elegans* reicht von Nukleotid 2805 bis 3653 in der SEQ ID NO: 34.

Tabakpflanzen wurden co-transformiert mit dem Vektor pSUN3CeLPLAT und dem in DE 102 19 203 und SEQ ID NO: 19 beschriebenen Vektor enthaltend Gene kodierend für Δ -6-Desaturase, Δ -6-Elongase und Δ -5-Desaturase, wobei die Selektion transgener Pflanzen mit Kanamycin erfolgte.

Tabakpflanzen wurden außerdem transformiert mit dem Vektor pGPTV USP/OCS-1,2,3 PSE1(Pp)+D6-Des(Pt)+2AT (T06E8-1) [siehe SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36 und SEQ ID NO: 37].

Lein wurde mit dem Vektor pSUN3CeLPLAT transformiert. Die entstandenen transgenen Pflanzen wurden mit solchen transgenen Leinpflanzen gekreuzt, die bereits geringe Mengen an ARA und EPA aufgrund der funktionellen Genexpression von Δ -6-Desaturase, Δ -6-Elongase und Δ -5-Desaturase enthielten.

Weiterhin wurde Lein mit dem Vektor pGPTV LegB4-700 + T06E8.1 transformiert. Die entstandenen transgenen Pflanzen wurden mit solchen transgenen Leinpflanzen gekreuzt, die bereits geringe Mengen an ARA und EPA aufgrund der funktionellen Expression von Δ -6-Desaturase, Δ -6-Elongase und Δ -5-Desaturase enthielten.

Die Samen von transgenen Tabak- und Leinpflanzen wurden wie weiter vorne beschrieben [Beispiel 3 b)] auf erhöhte Gehalte an LCPUFAs in untersucht.

Aus den hier vorliegenden Arbeiten lässt sich die Funktion der Acyl-CoA:Lyso-phospholipid-Acyltransferase (LPLAT) wie in Figur 10 dargestellt ableiten. Der Biosynthese-Weg der LCPUFAs stellt sich damit wie folgt dar.

Desaturasasen katalysieren die Einführung von Doppelbindungen in lipidgekoppelte Fettsäuren (*sn*2-Acyl-Phosphatidylcholin), während die Elongasen exklusiv die Elongation Coenzym A-veresterter Fettsäuren (Acyl-CoAs) katalysieren. Nach diesem Mechanismus erfordert die alternierende Wirkung von Desaturasasen und Elongasen einen ständigen Austausch von Acyl-Substraten zwischen Phospholipiden und Acyl-CoA-Pool und somit die Existenz einer zusätzlichen Aktivität, die die Acyl-Substrate in die jeweils notwendige Substratform, d.h. Lipide (für Desaturasasen) oder CoA-Thioester (für Elongasen), überführt. Dieser Austausch zwischen Acyl-CoA Pool und Phospholipiden wird durch LCPUFA-spezifische LPLAT ermöglicht. Die Biosynthese von ARA (A) erfolgt analog zu EPA (B), mit dem Unterschied, dass bei EPA der Δ -6-Desaturierung eine Δ -15-Desaturierung vorgeschaltet ist, so dass α 18:3-PC als Substrat für die Δ -6-Desaturase fungiert. Die Biosynthese von DHA macht einen weiteren Austausch zwischen Phospholipiden und Acyl-CoA-Pool über LPLAT notwendig: 20:5 ^{Δ 5,8,11,14,17} wird vom Phospholipid- zum CoA-Pool transferiert und nach erfolgter Δ -5-Elongation wird 22:5 ^{Δ 7,10,13,16,19} vom CoA- zum Phospholipid-Pool transferiert und schließlich durch Δ -4-Desaturase zu DHA umgesetzt. Gleiches gilt für den Austausch im Biosyntheseweg unter Verwendung der Δ -8-Desaturase, der Δ -9-Elongase und der Δ -5-Desaturase.

Äquivalente

Der Fachmann erkennt oder kann viele Äquivalente der hier beschriebenen erfindungsgemäßen spezifischen Ausführungsformen feststellen, indem er lediglich Routineexperimente verwendet. Diese Äquivalente sollen von den Patentansprüchen umfasst sein.

Patentansprüche

1. Isolierte Nukleinsäuresequenzen, die für Polypeptide mit Acyl-CoA:Lyso-phospholipid-Acyltransferaseaktivität codieren, wobei die durch die Nuklein-säuresequenzen codierten Acyl-CoA:Lyso-phospholipid-Acyltransferasen spezi-fisch C₁₆-, C₁₈-, C₂₀- oder C₂₂-Fettsäuren mit mindestens einer Doppelbindung im Fettsäuremolekül umsetzen.
2. Isolierte Nukleinsäuresequenzen gemäß Anspruch 1 ausgewählt aus der Grup-pe:
 - a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Sequenz,
 - b) Nukleinsäuresequenzen, die sich als Ergebnis des degenerierten geneti-schen Codes von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 enthaltenden codierenden Sequenz ableiten lassen
 - c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 darge-stellten Aminosäuresequenz codieren und mindestens 40 % Homologie auf Aminosäureebene mit SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 aufweisen und eine Acyl-CoA:Lyso-phospholipid-Acyltransferaseaktivität aufweisen.
3. Isolierte Nukleinsäuresequenz nach Anspruch 1 oder 2, wobei die Sequenz aus einem Eukaryont stammt.
4. Aminosäuresequenz, die von einer isolierten Nukleinsäuresequenz nach einem der Ansprüche 1 bis 3 codiert wird.
5. Genkonstrukt, enthaltend eine isolierte Nukleinsäure nach einem der Ansprüche 1 bis 3, wobei die Nukleinsäure funktionsfähig mit einem oder mehreren Regula-tionssignalen verbunden ist.
6. Genkonstrukt nach Anspruch 5, dadurch gekennzeichnet, dass das Nukleinsäu-rekonstrukt zusätzliche Biosynthesegene des Fettsäure- oder Lipidstoffwechsels enthält ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n).

7. Genkonstrukt nach Anspruch 5 oder 6, dadurch gekennzeichnet, dass das Nukleinsäurekonstrukt zusätzliche Biosynthesegene des Fettsäure- oder Lipidstoffwechsels enthält ausgewählt aus der Gruppe der Δ -4-Desaturase-, Δ -5-Desaturase-, Δ -6-Desaturase-, Δ -8-Desaturase-, Δ -9-Desaturase-, Δ -12-Desaturase-, Δ -5-Elongase-, Δ -6-Elongase- oder Δ -9-Elongase.
8. Vektor, enthaltend eine Nukleinsäure nach den Ansprüchen 1 bis 3 oder ein Genkonstrukt nach den Ansprüchen 5 bis 7.
9. Transgener nicht-humaner Organismus, enthaltend mindestens eine Nukleinsäure nach den Ansprüchen 1 bis 3, ein Genkonstrukt nach den Ansprüchen 5 bis 7 oder einen Vektor nach Anspruch 8.
10. Transgener nicht-humaner Organismus nach Anspruch 9, wobei der Organismus ein Mikroorganismus, ein nicht-humanes Tier oder eine Pflanze ist.
11. Transgener nicht-humaner Organismus nach Anspruch 9 oder 10, wobei der Organismus eine Pflanze ist.
12. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren in einem Organismus, dadurch gekennzeichnet, dass das Verfahren folgende Schritte umfasst:
- a) Einbringen mindestens einer Nukleinsäuresequenz in den Organismus mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Sequenz, die für ein Polypeptid mit einer Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität codiert; oder
 - b) Einbringen mindestens einer Nukleinsäuresequenz in den Organismus, die sich als Ergebnis des degenerierten genetischen Codes von der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 enthaltenden codierenden Sequenz ableiten lässt, oder
 - c) Einbringen mindestens eines Derivates der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 7 dargestellten Nukleinsäuresequenz in den Organismus, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 dargestellten Aminosäuresequenz codieren und mindestens 40 % Homologie auf Aminosäureebene mit SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 8 aufweisen und eine äquivalente Acyl-CoA:Lysophospholipid-Acyltransferaseaktivität aufweisen, und
 - d) kultivieren und ernten des Organismus.
13. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach Anspruch 12, dadurch gekennzeichnet, dass zusätzlich zu den unter (a) bis (c) genannten Nukleinsäuresequenzen weitere Nukleinsäuresequenzen in den Organismus eingebracht wurden, die für Polypeptide des Fettsäure- oder Lipidstoffwechsels

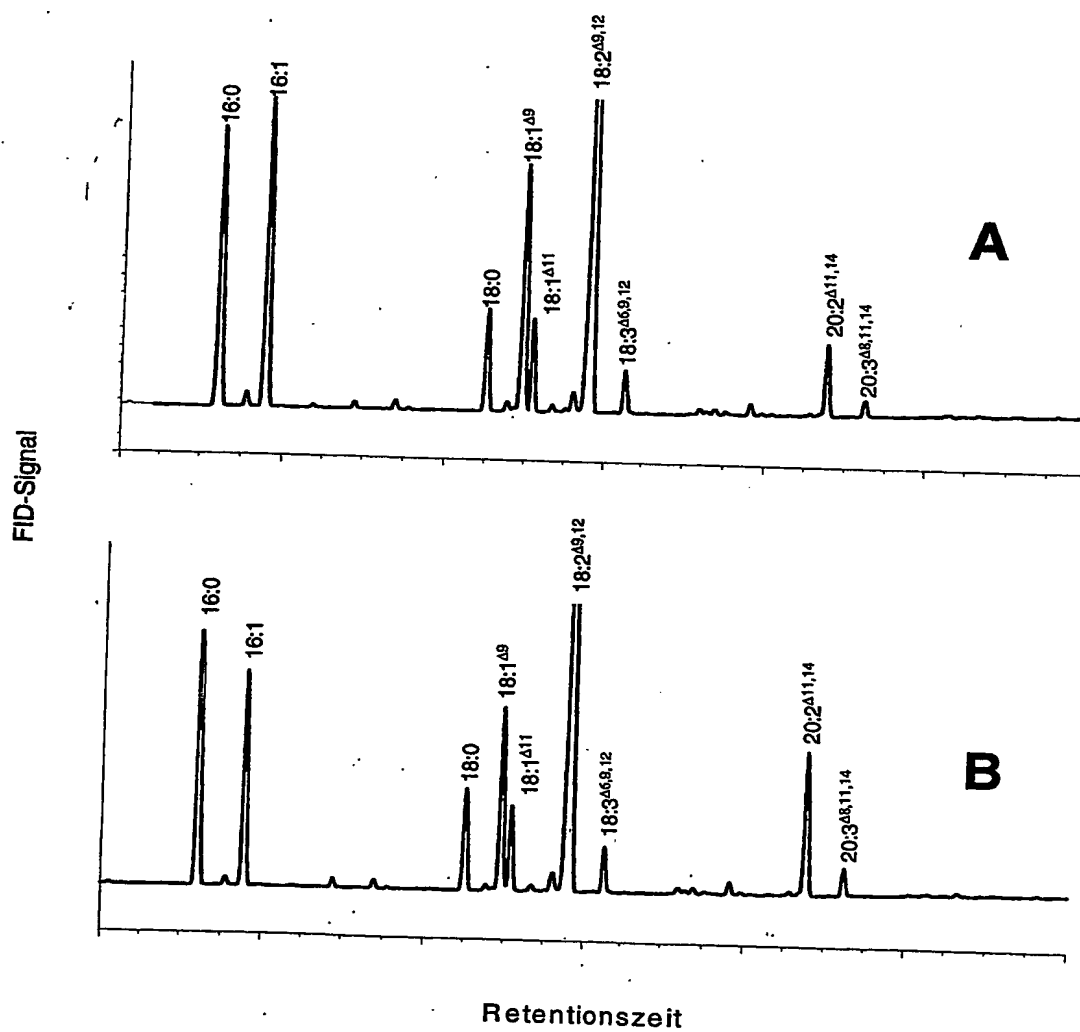
- ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) codieren.
5. 14. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach Anspruch 12 oder 13, dadurch gekennzeichnet, dass zusätzlich zu den unter (a) bis (c) genannten Nukleinsäuresequenzen weitere Nukleinsäuresequenzen in den Organismus eingebracht wurden, die für Polypeptide ausgewählt aus der Gruppe Δ -4-Desaturase-, Δ -5-Desaturase-, Δ -6-Desaturase-, Δ -8-Desaturase-, Δ -9-Desaturase-, Δ -12-Desaturase-, Δ -5-Elongase-, Δ -6-Elongase- oder Δ -9-Elongaseaktivität codieren.
- 10 15. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 14, dadurch gekennzeichnet, dass als Substrat der Acyl-CoA:Lyso-phospholipid-Acyltransferasen C_{16} -, C_{18} -, C_{20} - oder C_{22} -Fettsäuren verwendet werden.
- 15 16. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 15, dadurch gekennzeichnet, dass die mehrfach ungesättigten Fettsäuren aus dem Organismus in Form eines Öls, Lipids oder einer freien Fettsäure isoliert werden.
- 20 17. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 16, dadurch gekennzeichnet, dass die im Verfahren hergestellte mehrfach ungesättigte Fettsäure eine C_{18} -, C_{20} - oder C_{22} -Fettsäuren mit mindestens zwei Doppelbindungen im Molekül ist.
- 25 18. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 17, dadurch gekennzeichnet, dass im Verfahren eine mehrfach ungesättigte Fettsäure ausgewählt aus der Gruppe Dihomo- γ -linolensäure, Arachidonsäure, Eicosapentaensäure, Docosapentaensäure und Docosahexaensäure hergestellten wird.
- 30 19. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 18, dadurch gekennzeichnet, dass der Organismus ein Mikroorganismus, ein nicht-humanes Tier oder eine Pflanze ist.
- 35 20. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 19, dadurch gekennzeichnet, dass der Organismus eine transgene Pflanze ist.

21. Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren nach den Ansprüchen 12 bis 20, dadurch gekennzeichnet, dass die transgene Pflanze eine Ölfruchtpflanze ist.
- 5 22. Öl, Lipide oder Fettsäuren oder eine Fraktion davon, hergestellt durch das Verfahren nach einem der Ansprüche 12 bis 21.
23. Öl-, Lipid- oder Fettsäurezusammensetzung, die mehrfach ungesättigter Fettsäuren hergestellt nach einem Verfahren nach einem der Ansprüche 12 bis 21 umfasst und von transgenen Pflanzen stammt.
- 10 24. Verwendung von Öl, Lipide oder Fettsäuren hergestellt nach einem Verfahren nach einem der Ansprüche 12 bis 21 oder Öl-, Lipid- oder Fettsäurezusammensetzung gemäß Anspruch 23 in Futter, Nahrungsmitteln, Kosmetika oder Pharmazeutika.

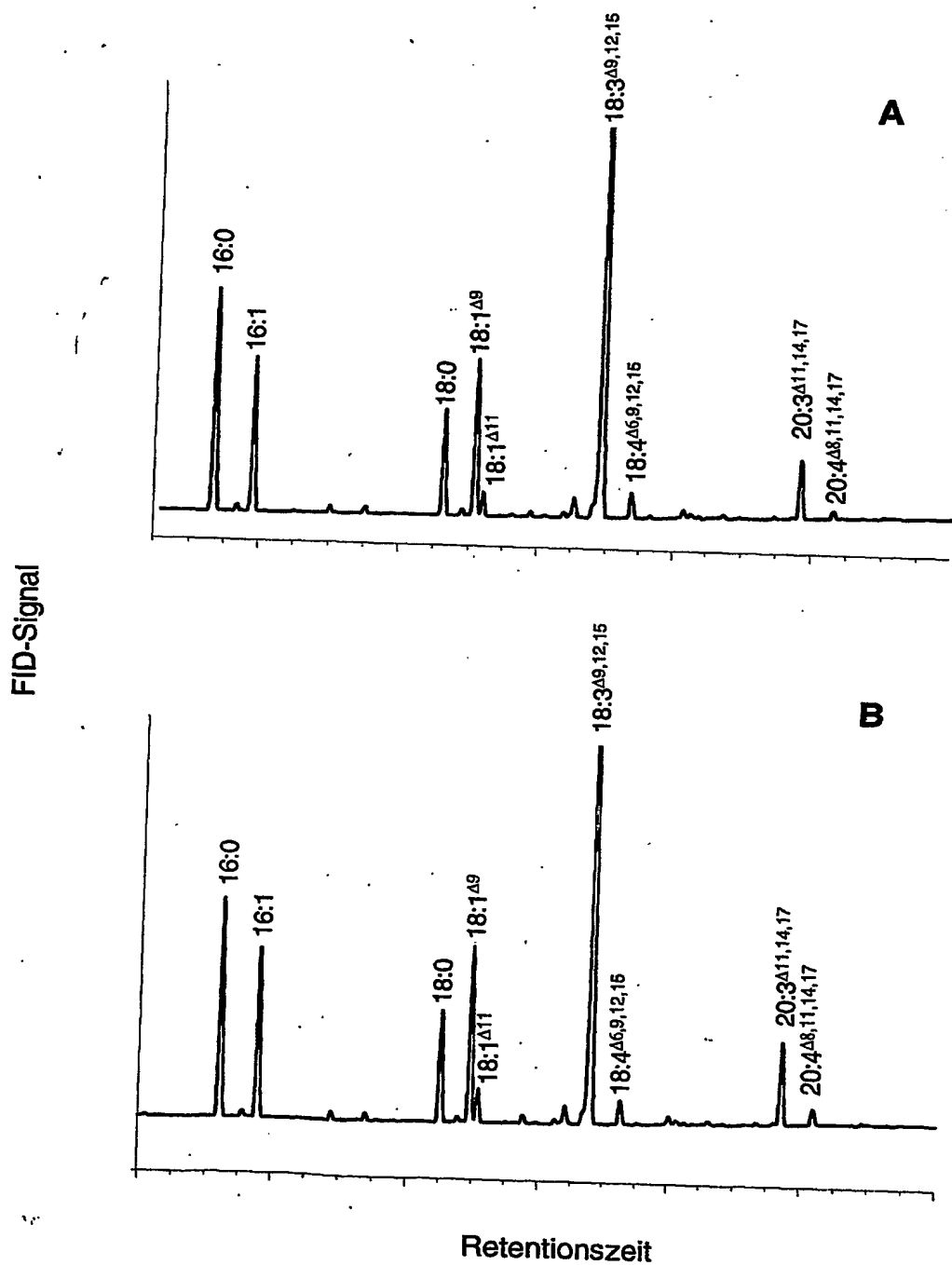
Figur 1: Aminosäure-Sequenzvergleich von *C. elegans* LPLATs (Ce-T06E8.1 und Ce-F59F4.4) mit der *M. musculus* LPAAT (Mm-NP061350).

	1				50
Mm-NP061350	MELWPGAWTA	LLLLLLLLLS	TLWFCSSSAK	YFFKMAFYNG	WELFLAILAI
Ce-T06E8.1	...MENFWSI	VVFFLSILF	ILYNISTVCH	YVMRISFYF	TILLHGMEVC
Ce-F59F4.4MTF	LAILEVIAVL	LLLAQLPVIG	FYIRAVYFGM	CLIGGFLGG
	51				100
Mm-NP061350	PVCAVRGRNV	ENMKIDRLLL	LHAKYLYCER	VEVRGAHHFP	PTQPYAAVSN
Ce-T06E8.1	VTMIPSWLNG	KGADYVFHSF	FYWCKWTCMH	TTVYGYEKTQ	VEGPAVVICN
Ce-F59F4.4	LASIPFGKSP	NNHFRMFKIF	QAMTWPMQVR	FELRNSEILH	DKKPYEELAN
	101				150
Mm-NP061350	HQSSLDLGLM	MEVLPDRCVP	IAKREILWAG	SAGLACWLAG	ITFIDRKRTG
Ce-T06E8.1	HQSSLDLGLM	ASIWPKNCVV	MMKRILAYVP	FFNLGAYFSN	ITFIDRYNRE
Ce-F59F4.4	HQSALDVLGM	SFAWPDVCVV	MLKSSLKXLP	GFNLCAVLCN	SVYINRFSKE
	151				200
Mm-NP061350	DAISVMSEVA	QTLILTQDVRV	WVFPECTRNH	NGSM PFKRC	AFHLAVQAQV
Ce-T06E8.1	RAMASVDYCA	SEMKNRNLIK	WVFPECTRNH	EGGFIPFKKG	AFNLAVRAQI
Ce-F59F4.4	KALKTMDTTL	HEIVTKKRKY	WVFPECTRNA	EPELIPFKKG	AFILAKQAKI
	201				250
Mm-NP061350	PIIPIVMSSY	QDFYSKKERR	FTSPGRCQVR	VLPPVSTIEGL	TPDDVPALAD
Ce-T06E8.1	PIIPVVFSDY	RDFYSKPGRY	FKNDGEVVLR	VLDATPTKGL	TLDDVSELSD
Ce-F59F4.4	PIVPCVFSSH	KFFYSHAERK	LTS.GNCIED	ILPEVDSS..	KFDSIDDLISA
	251			285	
Mm-NP061350	SVRHSMILTIF	REITSDGLGG	GDCLKKPGGA	GEARL	
Ce-T06E8.1	MCRDVMLAAY	KEVTLEAQOR	NATRRGETKD	GKKSE	
Ce-F59F4.4	HCKIMQAHK	EKLDAEAAANL	NI.....	

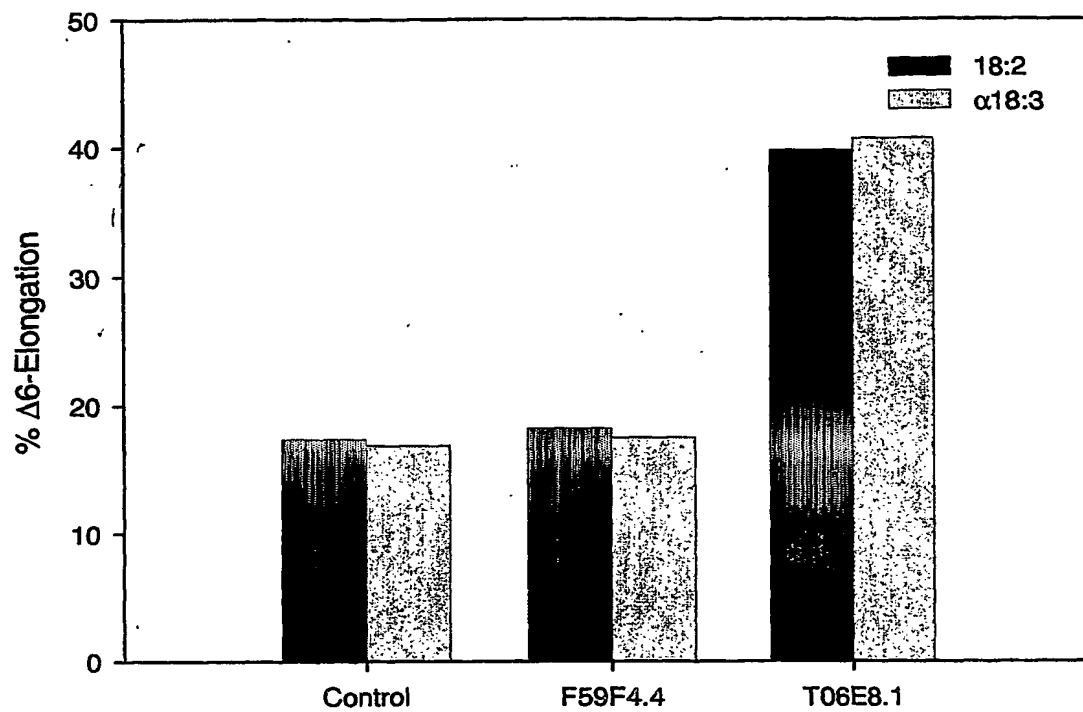
Figur.2: Fettsäureprofile von transgenen C13ABYS86 *S. cerevisiae*-Zellen



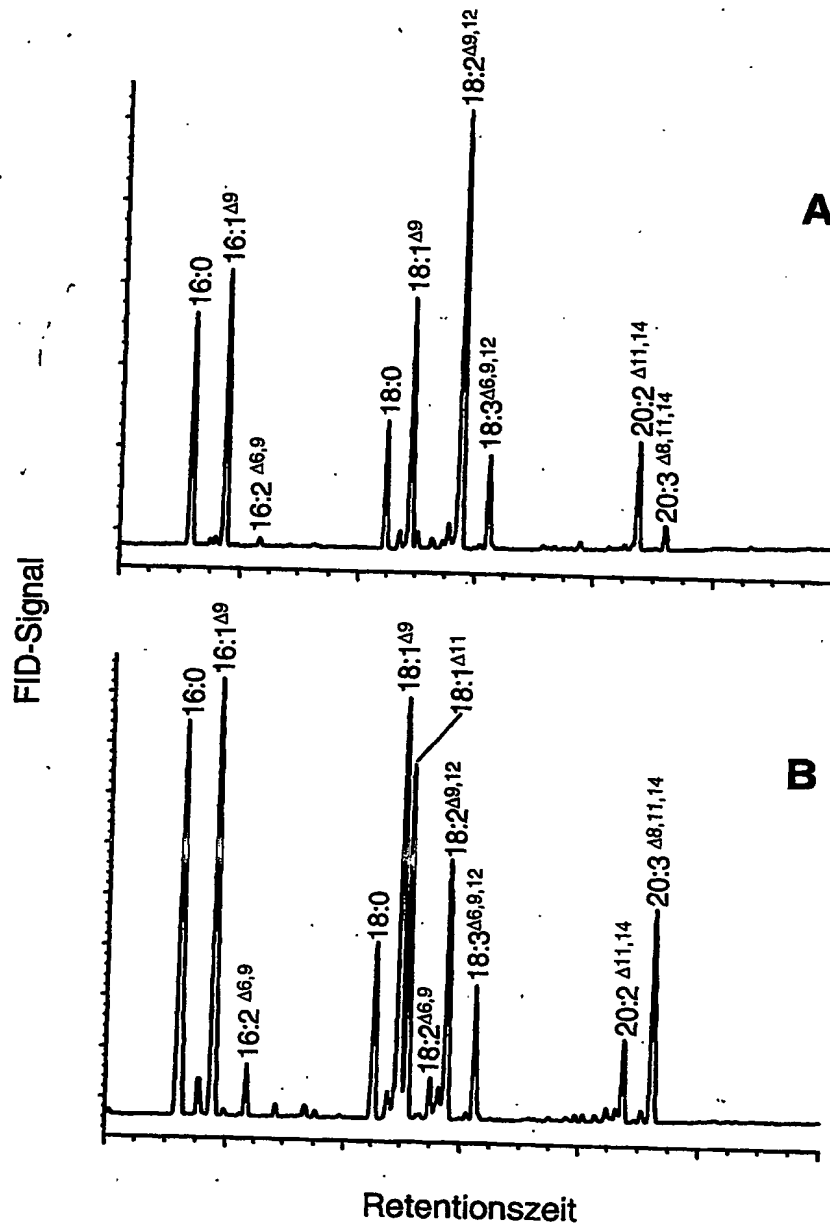
Figur 3: Fettsäureprofile von transgenen C13ABYS86 *S. cerevisiae*-Zellen



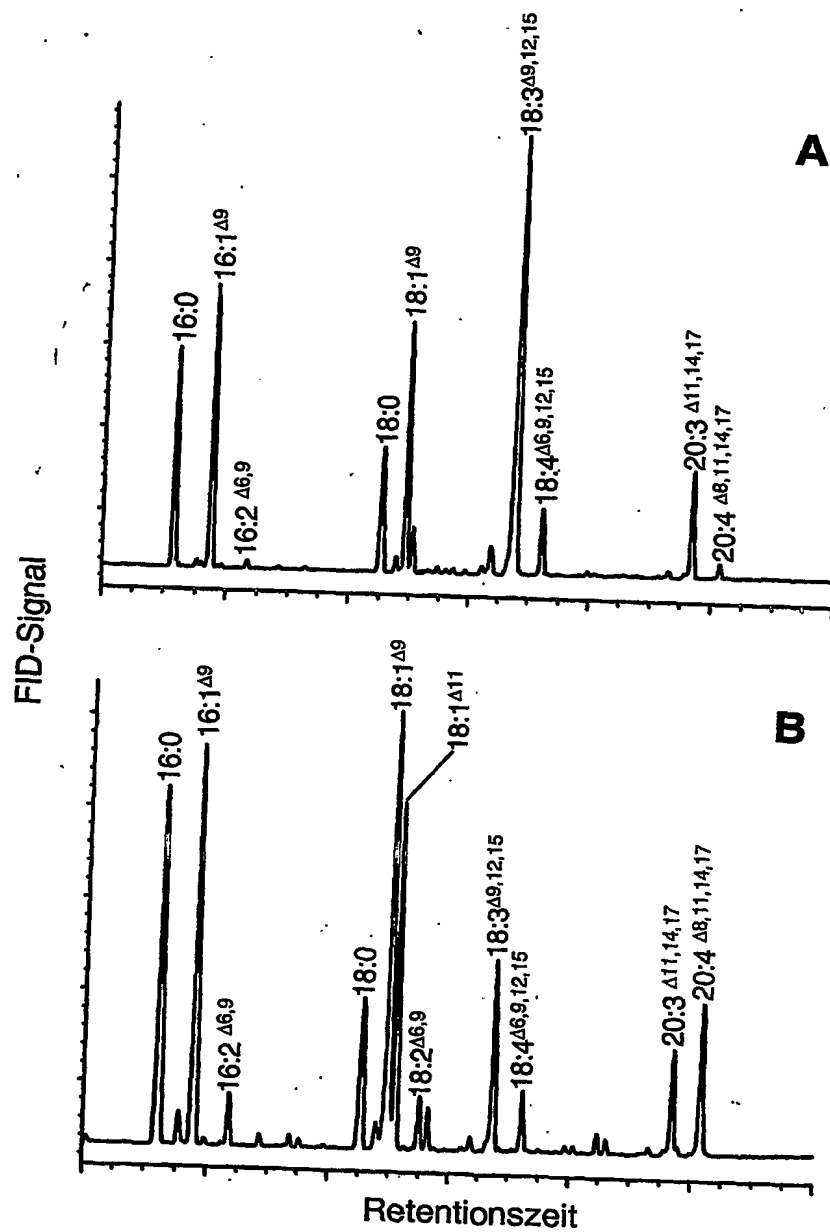
Figur 4: Elongation exogen applizierter $18:2^{\Delta 9,12}$ bzw. $18:3^{\Delta 9,12,15}$ im Anschluss an ihre endogene $\Delta 6$ -Desaturierung (Daten aus Fig. 2 und 3).



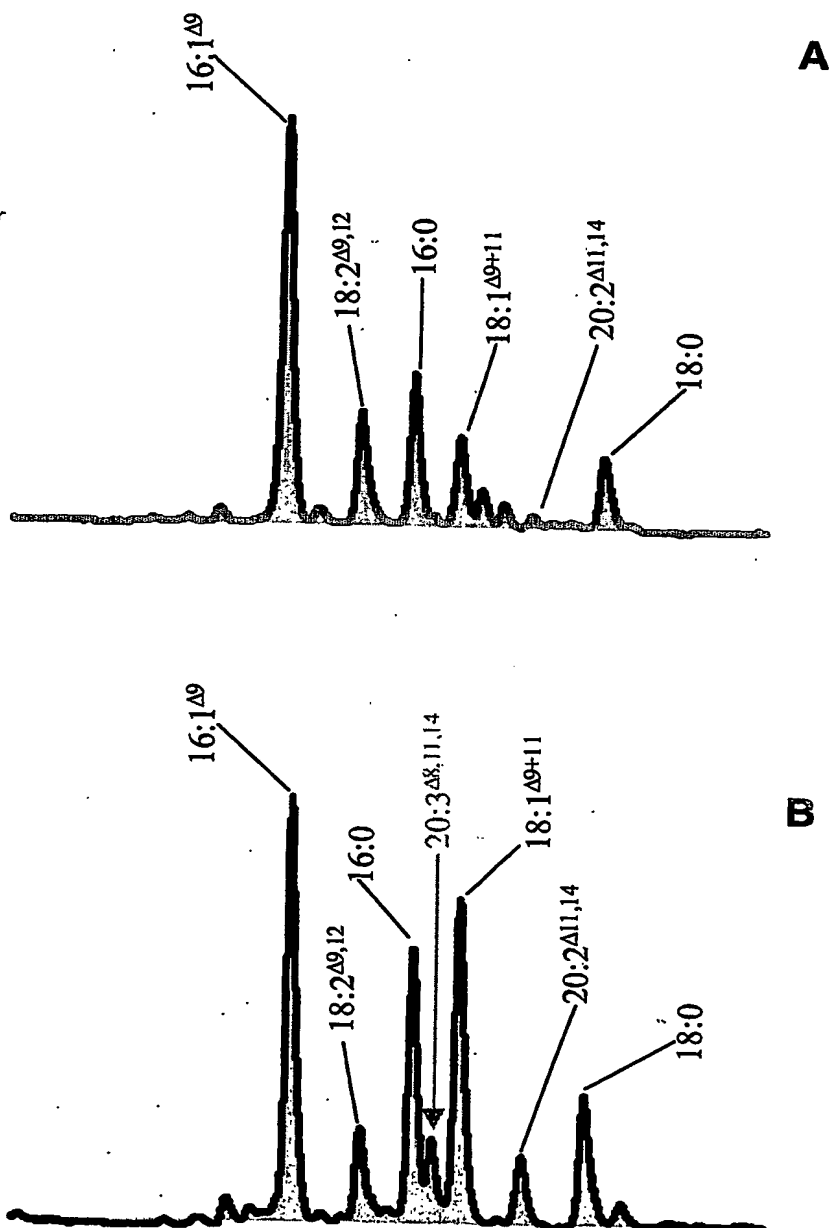
Figur 5: Fettsäure-Profile von transgenen INVSc1 *S. cerevisiae*-Zellen



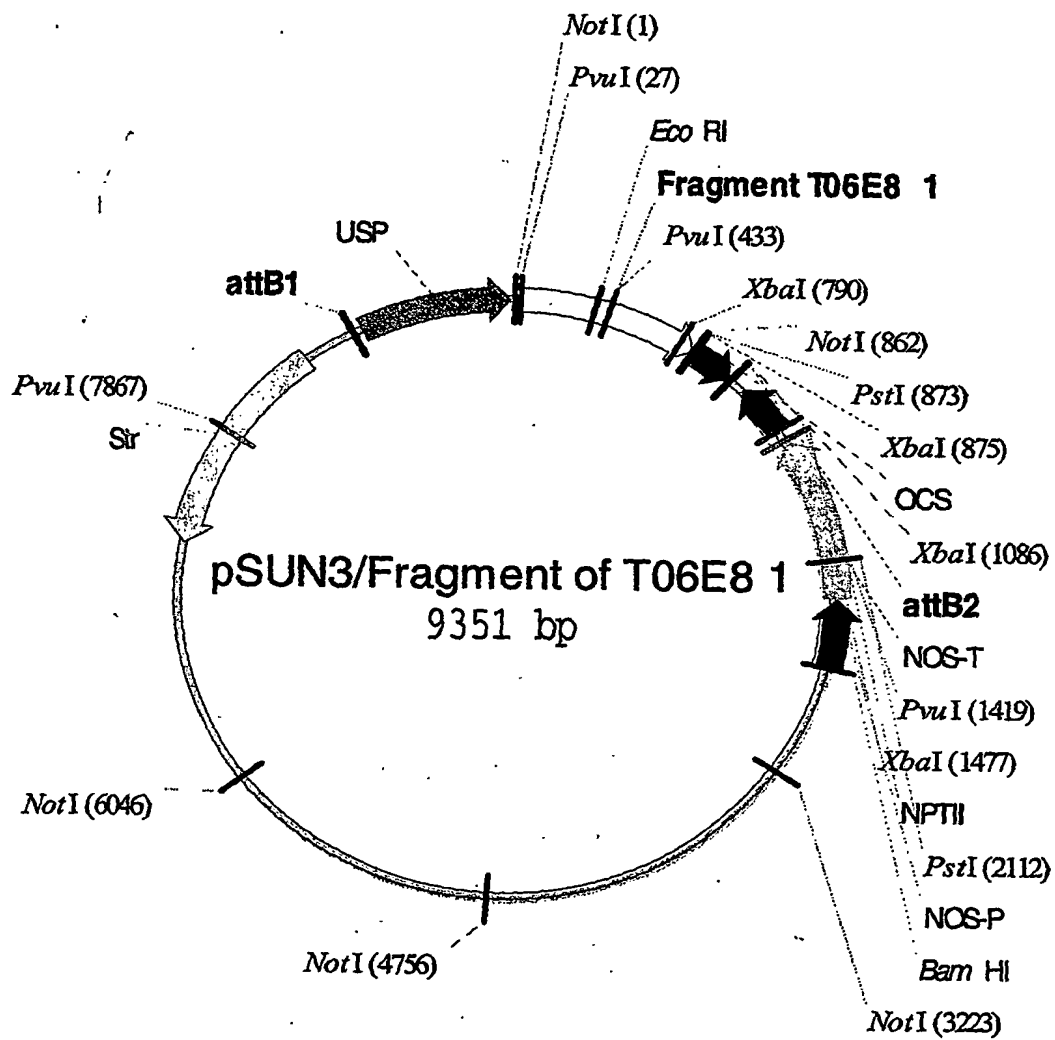
Figur 6: Fettsäure-Profile von transgenen INVSc1 *S. cerevisiae*-Zellen.



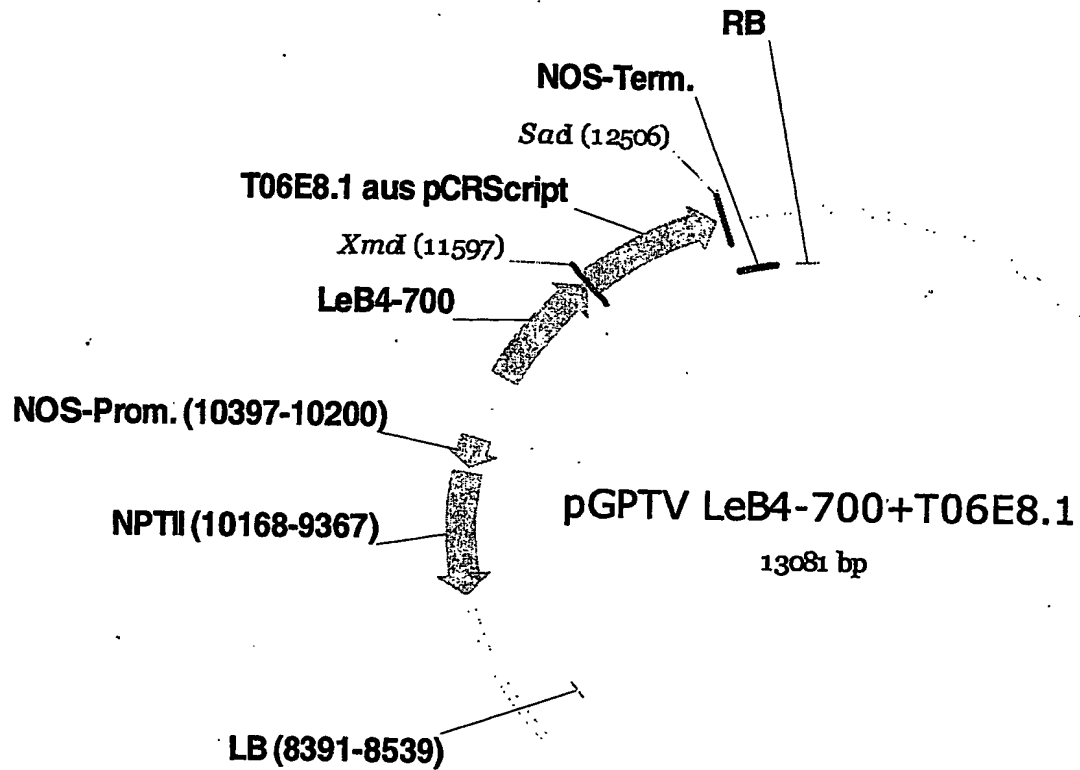
Figur 7: Acyl-CoA-Zusammensetzung transgener INVSc1 Hefen, die mit den Vektoren pESCLeu PpD6Pse1/pYes2 (A) oder pESCLeu-PpD6-Pse1/pYes2-T06E8.1 (B) transformiert worden waren.



Figur 8: Vektorkarte von pSUN3CeLPLAT

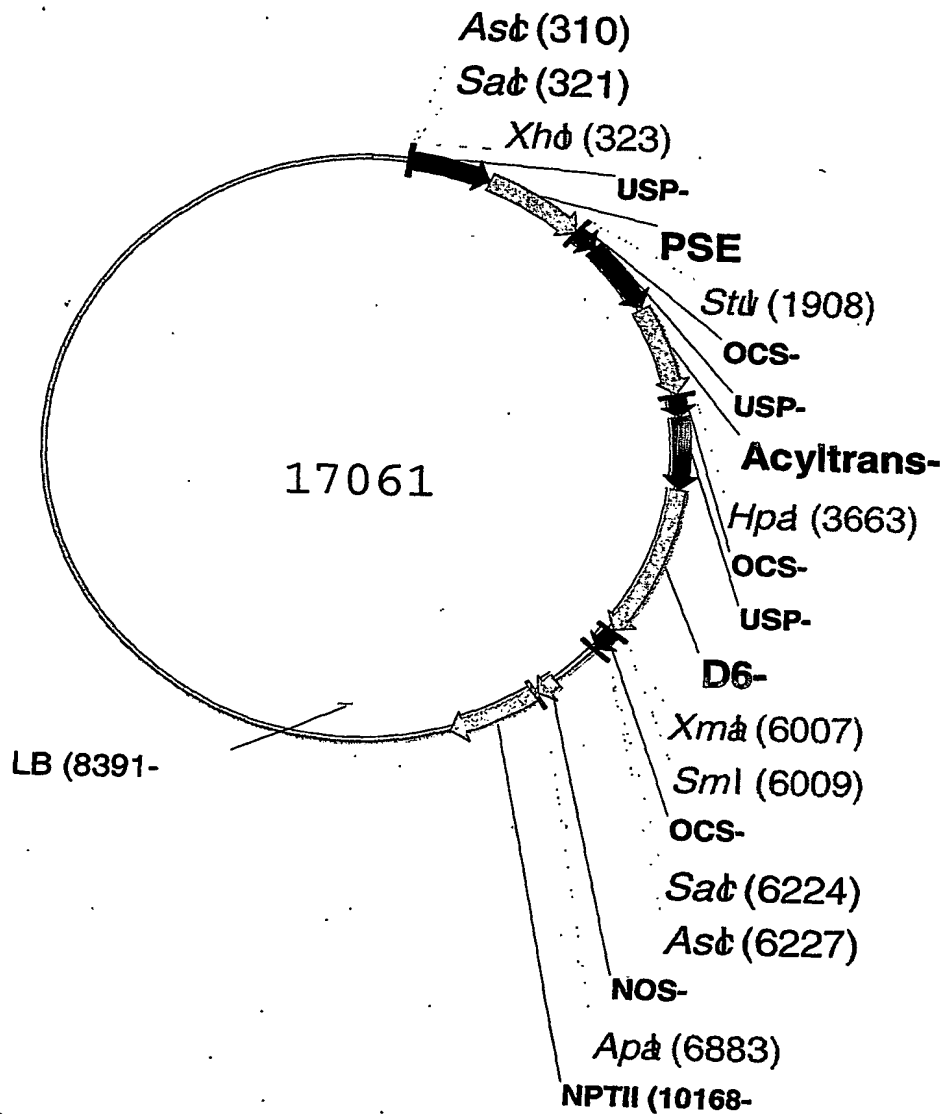


Figur 9A: Vektorkarte von pGPTV LeB4-700 + T06E8.1

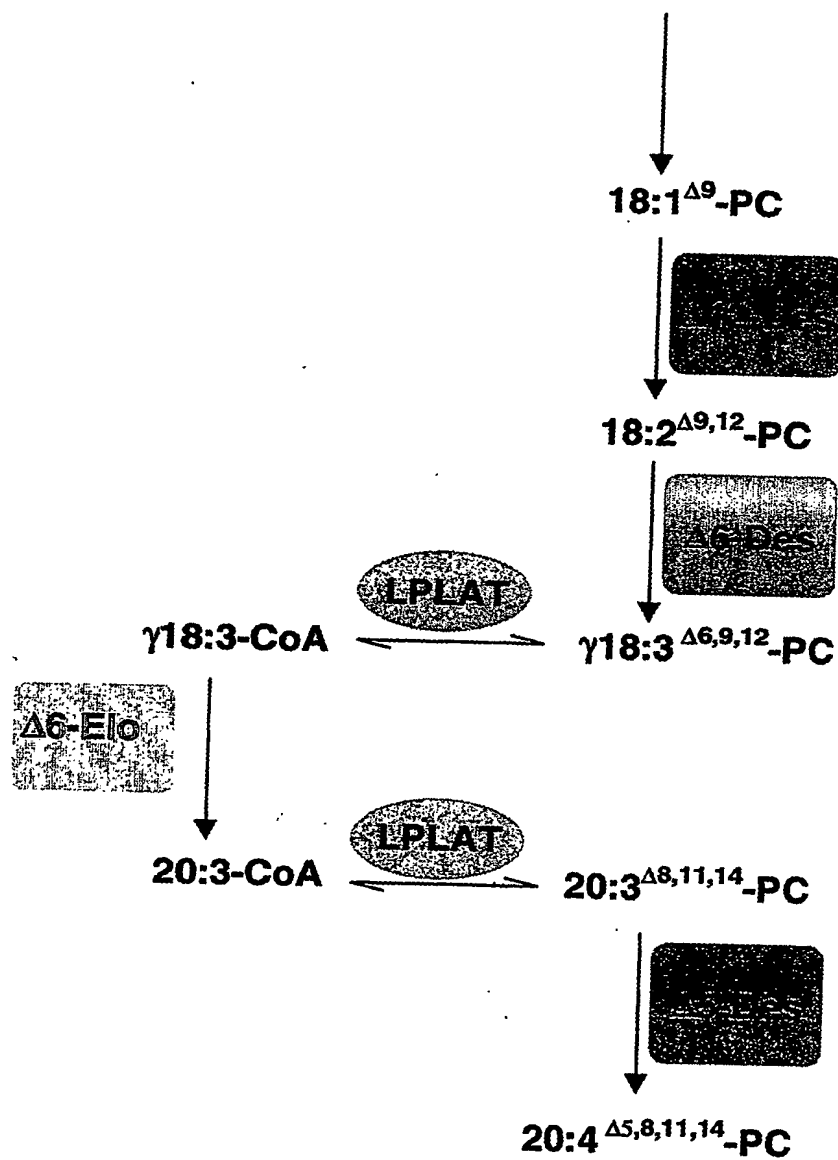


Figur 9B: Vektorkarte von pGPTV USP/OCS-1,2,3 PSE1(Pp)+D6-Des(Pt)+2AT (T06E8-1)

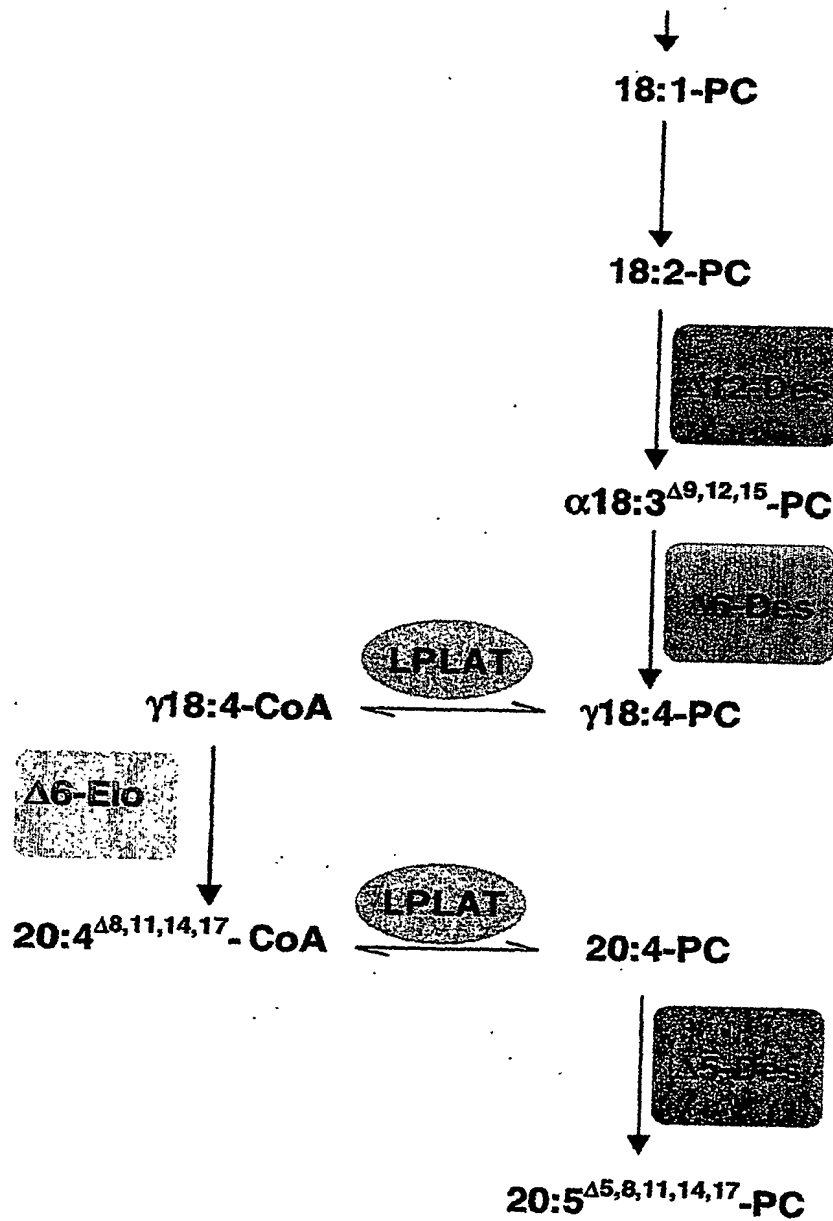
pGPTV/USP/OCS-1,2,3 PSE1 (Pp) D6-Des(Pt)-2 AT(T06E8-1)



Figur 10A: Biosynthese-Weg von LCPUFAs



Figur 10B: Biosynthese-Weg von LCPUFAs



SEQUENCE LISTING

<110> BASF Plant Science GmbH

5 <120> Verfahren zur Herstellung mehrfach ungesättigter Fettsäuren

<130> 20030015

10 <160> 37

<170> PatentIn version 3.1

<210> 1

<211> 849

15 <212> DNA

<213> *Caenorhabditis elegans*

<220>

<221> CDS

20 <222> (1) .. (849)

<223> Acyl-CoA:Lysophospholipid-Acyltransferase

<400> 1

25 atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc 48
Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1 5 10 15

30 ttc att tta tat aac ata tcg aca gta tgc cac tac tat atg cgg att 96
Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
20 25 30

35 tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt 144
Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
35 40 45

40 aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt 192
Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
50 55 60

45 cac tcg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc 240
His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
65 70 75 80

50 tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt 288
Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
85 90 95

55 aat cat cag agt tct ctc gac att cta tcg atg gca tca atc tgg ccg 336
Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
100 105 110

60 aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc 384
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
115 120 125

65 ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat 432
Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
130 135 140

aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg 480
Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met

	145	150	155	160	
5	aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn	165	170	175	528
10	cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala	180	185	190	576
	ggt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg	195	200	205	624
15	gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val	210	215	220	672
20	ggt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp	225	230	235	720
25	gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala	245	250	255	768
30	tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg	260	265	270	816
	gga gaa aca aaa gac ggg aag aaa tct gag taa Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu	275	280		849
35	<210> 2 <211> 282 <212> PRT <213> Caenorhabditis elegans				
40	<400> 2				
45	Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu	1	5	10	15
50	Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile	20	25	30	
55	Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val	35	40	45	
	Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe	50	55	60	
60	His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val	65	70	75	80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
 85 90 95
 5 Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110
 10 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125
 15 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140
 Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160
 20 Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
 165 170 175
 25 Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
 180 185 190
 30 Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
 195 200 205
 35 Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
 210 215 220
 Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
 225 230 235 240
 40 Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255
 45 Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
 260 265 270
 50 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
 275 280
 55 <210> 3
 <211> 849
 <212> DNA
 <213> Caenorhabditis elegans
 60 <220>
 <221> CDS
 <222> (1)..(849)
 <223> Acyl-CoA:Lysophospholipid-Acyltransferase

<400> 3		
5	atg gag aac ttc tgg tgg atc gtc gtg ttt ttt cta ctc tca att ctc Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu	48
	1 5 10 15	
10	ttc att tta tat aac ata tgg aca gta tgc cac tac tat atg cgg att Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile	96
	20 25 30	
15	tgg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val	144
	35 40 45	
20	aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe	192
	50 55 60	
25	cac tgg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val	240
	65 70 75 80	
30	tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys	288
	85 90 95	
35	aat gat cag agt tct ctc gac att cta tgg atg gca tca atc tgg ccg Asn His Gln Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro	336
	100 105 110	
40	aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe	384
	115 120 125	
45	ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr	432
	130 135 140	
50	aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met	480
	145 150 155 160	
55	aag aac aga aat ctt aaa ctt tgg gta tct ccg gaa gga aca aga aat Lys Asn Arg Asn Leu Lys Leu Trp Val Ser Pro Glu Gly Thr Arg Asn	528
	165 170 175	
60	cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala	576
	180 185 190	
65	gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg	624
	195 200 205	
70	gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val	672
	210 215 220	
75	gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp	720
	225 230 235 240	

gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc 768
 Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255

5 tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt 816
 Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
 260 265 270

10 gga gaa aca aaa gac ggg aag aaa tct gag taa 849
 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
 275 280

15 <210> 4
 <211> 282
 <212> PRT
 <213> Caenorhabditis elegans

20 <400> 4
 Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
 1 5 10 15

25 Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
 20 25 30

30 Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
 35 40 45

35 Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
 50 55 60

40 His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
 65 70 75 80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
 85 90 95

45 Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110

50 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125

55 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140

Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160

60 Lys Asn Arg Asn Leu Lys Leu Trp Val Ser Pro Glu Gly Thr Arg Asn
 165 170 175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
 180 185 190
 5
 Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
 195 200 205
 10
 Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
 210 215 220
 15
 Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
 225 230 235 240
 20
 Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255
 25
 Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
 260 265 270
 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
 275 280
 30
 <210> 5
 <211> 849
 <212> DNA
 <213> Caenorhabditis elegans
 35
 <220>
 <221> CDS
 <222> (1)..(849)
 <223> Acyl-CoA:Lysophospholipid-Acyltransferase
 40
 <400> 5
 atg gag aac ttc tgg tgc atc gtc gtg ttt ttt cta ctc tca att ctc 48
 Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
 1 5 10 15
 45
 ttc att tta tat aac ata tgc aca gta tgc cac tac tat gtg cgg att 96
 Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Val Arg Ile
 20 25 30
 50
 tgc ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt 144
 Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
 35 40 45
 55
 aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt 192
 Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
 50 55 60
 60
 cac tgc ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc 240
 His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
 65 70 75 80

	tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt	288
	Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys	
	85 90 95	
5	aat cat cag agt tct ctc gac att cta tcg atg gca tca atc tgg ccg	336
	Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro	
	100 105 110	
10	aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc	384
	Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe	
	115 120 125	
15	ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat	432
	Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr	
	130 135 140	
20	aac cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg	480
	Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met	
	145 150 155 160	
	aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat	528
	Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn	
	165 170 175	
25	cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca	576
	Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala	
	180 185 190	
30	gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg	624
	Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg	
	195 200 205	
35	gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt	672
	Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val	
	210 215 220	
40	gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat	720
	Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp	
	225 230 235 240	
	gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc	768
	Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala	
	245 250 255	
45	tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt	816
	Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg	
	260 265 270	
50	gga gaa aca aaa gac ggg aag aaa tct gag taa	849
	Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu	
	275 280	
55	<210> 6	
	<211> 282	
	<212> PRT	
	<213> Caenorhabditis elegans	
60	<400> 6	
	Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu	
	1 5 10 15	

5 Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Val Arg Ile
 20 25 30
 10 Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
 50 55 60
 15 His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
 65 70 75 80
 20 Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
 85 90 95
 25 Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110
 30 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125
 35 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140
 40 Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160
 45 Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
 165 170 175
 50 Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
 180 185 190
 55 Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
 195 200 205
 60 Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
 210 215 220
 Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
 225 230 235 240
 Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255

Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
 260 265 270

5 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
 275 280

10 <210> 7
 <211> 849
 <212> DNA
 <213> Caenorhabditis elegans

15 <220>
 <221> CDS
 <222> (1)..(849)
 <223> Acyl-CoA:Lysophospholipid-Acyltransferase

20 <400> 7
 atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc 48
 Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
 1 5 10 15

25 ttc att tta tat aac ata tcg aca gta tgc cac tac tat atg cgg att 96
 Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
 20 25 30

30 tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt 144
 Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
 35 40 45

35 aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt 192
 Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
 50 55 60

40 cac tcg ttt ttc tat tgg tgt aaa tgg act ggt gtt cat aca aca gtc 240
 His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
 65 70 75 80

tat gga tat gaa aaa aca caa gtt gaa ggt ccg gcc gta gtt att tgt 288
 Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
 85 90 95

45 aat cat cag ggt tct ctc gac att cta tcg atg gca tca atc tgg ccg 336
 Asn His Gln Gly Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110

50 aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc 384
 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125

55 ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat 432
 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140

60 aag cgt gaa cgt gcg atg gct tca gtt gat tat tgt gca tct gaa atg 480
 Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160

aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat 528
 Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn

	165	170	175		
5	cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att gca Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala	180	185	190	576
10	ggt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat cgg Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg	195	200	205	624
15	gat ttc tac tca aag cca ggc cga tat ttc aag aat gat gga gaa gtt Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val	210	215	220	672
20	gac gtc agc gag ttg tct gat atg tgt cgg gac gtt atg ttg gca gcc Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala	230	235	240	720
25	tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg	260	265	270	816
30	gga gaa aca aaa gac ggg aag aaa tct gag taa Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu	275	280		849
35	<210> 8 <211> 282 <212> PRT <213> Caenorhabditis elegans				
40	Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu	1	5	10	15
45	Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile	20	25	30	
50	Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val	35	40	45	
55	Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe	50	55	60	
60	His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val	65	70	75	80
	Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys	85	90	95	

Asn His Gln Gly Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110

5 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125

10 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140

15 Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160

20 Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
 165 170 175

Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
 180 185 190

25 Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
 195 200 205

30 Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
 210 215 220

35 Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
 225 230 235 240

40 Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255

Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
 260 265 270

45 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
 275 280

50 <210> 9
 <211> 1578
 <212> DNA
 <213> Physcomitrella patens

55 <220>
 <221> CDS
 <222> (1)..(1578)
 <223> Delta-6-Desaturase

60 <400> 9
 atg gta ttc gcg ggc ggt gga ctt cag cag ggc tct ctc gaa gaa aac
 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn

1	5	10	15	
atc gac gtc gag cac att gcc agt atg tct ctc ttc agc gac ttc ttc Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe	20	25	30	96
agt tât gtg tct tca act gtt ggt tcg tgg agc gta cac agt ata caa Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln	35	40	45	144
cct ttg aag cgc ctg acg agt aag aag cgt gtt tcg gaa agc gct gcc Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala	50	55	60	192
gtg caa tgt ata tca gct gaa gtt cag aga aat tcg agt acc cag gga Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly	65	70	75	240
act gcg gag gca ctc gca gaa tca gtc gtg aag ccc acg aga cga agg Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg	85	90	95	288
tca tct cag tgg aag aag tcg aca cac ccc cta tca gaa gta gca gta Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val	100	105	110	336
cac aac aag cca agc gat tgc tgg att gtt gta aaa aac aag gtg tat His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr	115	120	125	384
gat gtt tcc aat ttt gcg gac gag cat ccc gga gga tca gtt att agt Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser	130	135	140	432
act tat ttt gga cga gac ggc aca gat gtt ttc tct agt ttt cat gca Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala	145	150	155	480
gct tct aca tgg aaa att ctt caa gac ttt tac att ggt gac gtg gag Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu	165	170	175	528
agg gtg gag ccg act cca gag ctg ctg aaa gat ttc cga gaa atg aga Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg	180	185	190	576
gct ctt ttc ctg agg gag caa ctt ttc aaa agt tcg aaa ttg tac tat Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr	195	200	205	624
gtt atg aag ctg ctc acg aat gtt gct att ttt gct gcg agc att gca Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala	210	215	220	672
ata ata tgt tgg agc aag act att tca gcg gtt ttg gct tca gct tgt Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys	225	230	235	720
atg atg gct ctg tgt ttc caa cag tgc gga tgg cta tcc cat gat ttt Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe	245	250	255	768

	ctc cac aat cag gtg ttt gag aca cgc tgg ctt aat gaa gtt gtc ggg Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly	816
	260 265 270	
5	tat gtg atc ggc aac gcc gtt ctg ggg ttt agt aca ggg tgg tgg aag Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys	864
	275 280 285	
10	gag aag cat aac ctt cat cat gct gct cca aat gaa tgc gat cag act Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr	912
	290 295 300	
15	tac caa cca att gat gaa gat att gat act ctc ccc ctc att gcc tgg Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp	960
	305 310 315 320	
20	agc aag gac ata ctg gcc aca gtt gag aat aag aca ttc ttg cga atc Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile	1008
	325 330 335	
	ctc caa tac cag cat ctg ttc ttc atg ggt ctg tta ttt ttc gcc cgt Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg	1056
	340 345 350	
25	ggt agt tgg ctc ttt tgg agc tgg aga tat acc tct aca gca gtg ctc Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu	1104
	355 360 365	
30	tca cct gtc gac agg ttg ttg gag aag gga act gtt ctg ttt cac tac Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr	1152
	370 375 380	
35	ttt tgg ttc gtc ggg aca gcg tgc tat ctt ctc cct ggt tgg aag cca Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro	1200
	385 390 395 400	
40	tta gta tgg atg gcg gtg act gag ctc atg tcc ggc atg ctg ctg ggc Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly	1248
	405 410 415	
	ttt gta ttt gta ctt agc cac aat ggg atg gag gtt tat aat tcg tct Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser	1296
	420 425 430	
45	aaa gaa ttc gtg agt gca cag atc gta tcc aca cgg gat atc aaa gga Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly	1344
	435 440 445	
50	aac ata ttc aac gac tgg ttc act ggt ggc ctt aac agg caa ata gag Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu	1392
	450 455 460	
55	cat cat ctt ttc cca aca atg ccc agg cat aat tta aac aaa ata gca His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala	1440
	465 470 475 480	
	cct aga gtg gag gtg ttc tgt aag aaa cac ggt ctg gtg tac gaa gac Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp	1488
	485 490 495	
60	gta tct att gct acc ggc act tgc aag gtt ttg aaa gca ttg aag gaa Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu	1536

500 505 510

gtc gcg gag gct gcg gca gag cag cat gct acc acc agt taa
 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser 1578
 515 520 525

<210> 10
 <211> 525
 10 <212> PRT
 <213> Physcomitrella patens
 <400> 10

15 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
 1 5 10 15

20 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
 20 25 30

25 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
 35 40 45

30 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
 50 55 60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly
 65 70 75 80

35 Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg
 85 90 95

40 Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
 100 105 110

45 His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
 115 120 125

50 Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
 130 135 140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
 145 150 155 160

55 Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
 165 170 175

60 Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg
 180 185 190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr
 195 200 205

5 Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala
 210 215 220

10 Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys
 225 230 235 240

15 Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe
 245 250 255

20 Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly
 260 265 270

Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys
 275 280 285

25 Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr
 290 295 300

30 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp
 305 310 315 320

35 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile
 325 330 335

40 Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg
 340 345 350

Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu
 355 360 365

45 Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr
 370 375 380

50 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro
 385 390 395 400

55 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly
 405 410 415

Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser
 420 425 430

60 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
 435 440 445

5 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu
 450 455 460
 10 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
 485 490 495
 15 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
 500 505 510
 20 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
 515 520 525
 25 <210> 11
 <211> 1192
 <212> DNA
 <213> Physcomitrella patens
 30 <220>
 <221> CDS
 <222> (58)..(930)
 <223> Delta-6-Elongase
 35 <400> 11
 ctgcttcgtc tcattcttggg ggtgtgattc gggagtgggt tgagttggtg gagcgca 57
 atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg 105
 Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
 1 5 10 15
 40 cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg gat 153
 Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
 20 25 30
 45 acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc atc 201
 Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
 35 40 45
 50 gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt ttg 249
 Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
 50 55 60
 55 tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt ttg 297
 Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
 65 70 75 80
 60 ctg caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc agt 345
 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95
 ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg tac 393
 Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr

	100	105	110	
5	tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg att Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile 115 120 125			441
10	ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat acc Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr 130 135 140			489
	ggt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc cac Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His 145 150 155 160			537
15	ggt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct cat Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His 165 170 175			585
20	cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca gga His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly 180 185 190			633
25	gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt cga Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg 195 200 205			681
30	agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac ttg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu 210 215 220			729
	aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct tac Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr 225 230 235 240			777
35	tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile 245 250 255			825
40	ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt tac Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr 260 265 270			873
45	gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys 275 280 285			921
50	act gag tga gctgtatcaa gccatagaaa ctctattatg ttagaacctg Thr Glu 290			970
	aagttggtgc tttcttatct ccacttatct tttaagcagc atcagttttg aaatgatgtg			1030
	tgggcgtggt ctgcaagtag tcatcaatat aatcggcctg agcacttcag atggattgtt			1090
55	agaacatgag taaaagcggg tattacgggtg tttattttgt accaaatcac cgcacgggtg			1150
	aattgaaata tttcagattt gatcaatttc atctgaaaaa aa			1192
60	<210> 12 <211> 290 <212> PRT			

<213> Physcomitrella patens

<400> 12

5 Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
 1 5 10 15
 10 Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
 20 25 30
 15 Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
 35 40 45
 20 Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
 50 55 60
 25 Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
 65 70 75 80
 30 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95
 35 Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
 115 120 125
 40 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
 130 135 140
 45 Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
 145 150 155 160
 50 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
 180 185 190
 55 Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
 195 200 205
 60 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
 210 215 220
 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
 225 230 235 240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
 245 250 255
 5
 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
 260 265 270
 10
 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
 275 280 285
 15 Thr Glu
 290
 20
 <210> 13
 <211> 1410
 <212> DNA
 <213> Phaeodactylum tricornutum
 25
 <220>
 <221> CDS
 <222> (1)..(1410)
 <223> Delta-5-Desaturase
 30
 <400> 13
 atg gct ccg gat gcg gat aag ctt cga caa cgc cag acg act gcg gta 48
 Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
 1 5 10 15
 35
 gcg aag cac aat gct gct acc ata tcg acg cag gaa cgc ctt tgc agt 96
 Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
 20 25 30
 40
 ctg tct tcg ctc aaa ggc gaa gaa gtc tgc atc gac gga atc atc tat 144
 Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
 35 40 45
 45
 gac ctc caa tca ttc gat cat ccc ggg ggt gaa acg atc aaa atg ttt 192
 Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe
 50 55 60
 50
 ggt ggc aac gat gtc act gta cag tac aag atg att cac ccg tac cat 240
 Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His
 65 70 75 80
 55
 acc gag aag cat ttg gaa aag atg aag cgt gtc ggc aag gtg acg gat 288
 Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp
 85 90 95
 60
 ttc gtc tgc gag tac aag ttc gat acc gaa ttt gaa cgc gaa atc aaa 336
 Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys
 100 105 110
 60
 cga gaa gtc ttc aag att gtg cga cga ggc aag gat ttc ggt act ttg 384
 Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu
 115 120 125

	gga tgg ttc ttc cgt gcg ttt tgc tac att gcc att ttc ttc tac ctg Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu 130 135 140	432
5	cag tac cat tgg gtc acc acg gga acc tct tgg ctg ctg gcc gtg gcc Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala 145 150 155 160	480
10	tac gga atc tcc caa gcg atg att ggc atg aat gtc cag cac gat gcc Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala 165 170 175	528
15	aac cac ggg gcc acc tcc aag cgt ccc tgg gtc aac gac atg cta gcc Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly 180 185 190	576
20	ctc ggt gcg gat ttt att ggt ggt tcc aag tgg ctc tgg cag gaa caa Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln 195 200 205	624
	cac tgg acc cac cac gct tac acc aat cac gcc gag atg gat ccc gat His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp 210 215 220	672
25	agc ttt ggt gcc gaa cca atg ctc cta ttc aac gac tat ccc ttg gat Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp 225 230 235 240	720
30	cat ccc gct cgt acc tgg cta cat cgc ttt caa gca ttc ttt tac atg His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met 245 250 255	768
35	ccc gtc ttg gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile 260 265 270	816
40	ctt gac ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp 275 280 285	864
	aac gct ttc att cac tcg cga cgc aag tat gcg gtt ttc tgg cgg gct Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala 290 295 300	912
45	gtg tac att gcg gtg aac gtg att gct ccg ttt tac aca aac tcc gcc Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly 305 310 315 320	960
50	ctc gaa tgg tcc tgg cgt gtc ttt gga aac atc atg ctc atg ggt gtg Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val 325 330 335	1008
55	gcg gaa tcg ctc gcg ctg gcg gtc ctg ttt tcg ttg tcg cac aat ttc Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe 340 345 350	1056
60	gaa tcc gcg gat cgc gat ccg acc gcc cca ctg aaa aag acg gga gaa Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu 355 360 365	1104
	cca gtc gac tgg ttc aag aca cag gtc gaa act tcc tgc act tac ggt Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly	1152

	370	375	380	
5	gga ttc ctt tcc ggt tgc ttc acg gga ggt ctc aac ttt cag gtt gaa Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu 385 390 395 400			1200
10	cac cac ttg ttc cca cgc atg agc agc gct tgg tat ccc tac att gcc His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala 405 410 415			1248
15	ccc aag gtc cgc gaa att tgc gcc aaa cac ggc gtc cac tac gcc tac Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr 420 425 430			1296
20	tac ccg tgg atc cac caa aac ttt ctc tcc acc gtc cgc tac atg cac Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His 435 440 445			1344
25	gcg gcc ggg acc ggt gcc aac tgg cgc cag atg gcc aga gaa aat ccc Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro 450 455 460			1392
30	ttg acg gga cgg gcg taa Leu Thr Gly Arg Ala 465			1410
	<210> 14 <211> 469 <212> PRT <213> Phaeodactylum tricornutum <400> 14			
35	Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val 1 5 10 15			
40	Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser 20 25 30			
45	Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr 35 40 45			
50	Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe 50 55 60			
55	Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His 65 70 75 80			
60	Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp 85 90 95			
	Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys 100 105 110			

Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu
 115 120 125

5 Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu
 130 135 140

10 Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala
 145 150 155 160

15 Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala
 165 170 175

20 Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly
 180 185 190

25 Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln
 195 200 205

30 His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp
 210 215 220

35 Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp
 225 230 235 240

40 His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met
 245 250 255

45 Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile
 260 265 270

50 Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp
 275 280 285

55 Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala
 290 295 300

60 Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly
 305 310 315 320

65 Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val
 325 330 335

70 Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe
 340 345 350

75 Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu
 355 360 365

	tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc	660
	tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgaga tttaattgtt	720
5	gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg	780
	taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca	840
10	agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt	900
	ttaaaaatat ttgggaaatg atttgcacgtg aagccatgtg taaaaccatg acatccactt	960
	ggaggatgca ataatgaaga aaactacaaa ttacatgca actagttatg catgtagtct	1020
15	atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta	1080
	taattttctt atagccagcc caccgcgggtg ggcggccgcc tgcagtctag aaggcctcct	1140
20	gctttaatga gatatgcgag acgcctatga tcgcatgata ttgcttttca attctgttgt	1200
	gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgccgggt ttcggttcat	1260
	tctaataaat atatcacccg ttactatcgt atttttatga ataataattct ccgttcaatt	1320
25	tactgattgt ccgtcgacga attcgagctc ggcgcgccaa gcttggcgta atcatggtca	1380
	tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga	1440
30	agcataaagt gtaaaacctg ggggtgcctaa tgagtgaagt aactcacatt aattgcgttg	1500
	cgctcaactgc ccgctttcca gtccgggaaac ctgtcgtgcc agctgcatta atgaatcggc	1560
	caacgcgcgg ggagaggcgg tttgcgtatt ggcgcgtcct ccgcttcctc gctcaactgac	1620
35	tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcaactcaaa ggcggtaata	1680
	cggttatcca cagaatcagg ggataacgca ggaagaaca tgtgagcaaa aggccagcaa	1740
40	aaggccagga accgtaaaaa ggcgcggttg ctggcgtttt tccataggct ccgccccct	1800
	gacgagcatc aaaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa	1860
	agataccagg cgtttcccc tggaagctcc ctctgcgct ctctgttcc gaccctgccg	1920
45	cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca	1980
	cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa	2040
50	cccccgctt agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg	2100
	gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg	2160
	tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta cactagaagg	2220
55	acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc	2280
	tcttgatccg gcaaacaac caccgctggt agcgggtggt tttttgtttg caagcagcag	2340
60	attacgcgca gaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac	2400
	gctcagtgga acgaaaactc acgttaaggg attttgggtca tgagattatc aaaaaggatc	2460

	tccacctaga tcccttttaa ttaaaaatga agtttttaaat caatctaaag tatatatgag	2520
	taaacttggt ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt	2580
5	ctatttcgtt catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag	2640
	ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca	2700
10	gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcttgcaact	2760
	ttatccgcct ccattccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca	2820
	gttaatatgt tgcgcaacgt tggtgccatt gctacaggca tcgtgggtgtc acgctcgtcg	2880
15	tttgggatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc	2940
	atgttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg	3000
20	gccgcagtgt tatcactcat gggttatggca gactgcata attctcttac tgtcatgcca	3060
	tccgtaagat gcttttctgt gactgggtgag tactcaacca agtcattctg agaatagtgt	3120
	atgcggcgac cgagttgctc ttgccccggc tcaatacggg ataataccgc gccacatagc	3180
25	agaactttaa aagtgtcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc	3240
	ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg caccacaactg atcttcagca	3300
30	tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa	3360
	aagggaataa gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat	3420
	tgaagcattt atcaggggta ttgtctcatg agcggatata tatttgaaatg tatttagaaa	3480
35	aataaacaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa	3540
	accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc	3598
40	<210> 16	
	<211> 3590	
	<212> DNA	
	<213> artificial sequence	
45	<220>	
	<221> misc_feature	
	<223> Sequenz stellt eine pflanzliche Promotor-Terminator-Expressionska ssette in Vektor pUC19 dar	
50	<400> 16	
	tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca	60
55	cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg	120
	ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
	accaätatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
60	attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcggggc tcttcgctat	300
	tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360

	tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgcgcg agctcctcga	420
5	gcaaatttac acattgccac taaacgtcta aacccttgta atttggtttt gttttactat	480
	gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct	540
	tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta	600
10	tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc	660
	tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattggt	720
15	gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg	780
	taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaagggttag taatttttca	840
	agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt	900
20	ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt	960
	ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct	1020
25	atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta	1080
	taattttctt atagccagcg gatccgatat cgggcccgt agcgtaacc ctgctttaat	1140
	gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctggt gtgcacgttg	1200
30	taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attctaataga	1260
	atatatcacc cgttactatc gtattttttat gaataatatt ctccgttcaa tttactgatt	1320
35	gtccgctcgac gaattcgagc tcggcgcgcc aagcttgagg taatcatggt catagctggt	1380
	tcctgtgtga aattgttatc cgctcacaat tccacacaa atacgagccg gaagcataaa	1440
	gtgtaaagcc tgggggtgct aatgagttag ctaactcaca ttaattgcgt tgcgctcact	1500
40	gcccgttttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc	1560
	ggggagaggg ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgtgcg	1620
45	ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc	1680
	cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag	1740
	gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca	1800
50	tcacaaaaat cgacgctcaa gtcagagggt gcgaaacccg acaggactat aaagatacca	1860
	ggcgttttcc cctggaagct cctcgtgcg ctctcctggt ccgacctgc cgcttaccgg	1920
55	atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag	1980
	gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt	2040
	tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca	2100
60	cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg	2160
	cggtgctaca gatttcttga agtgggtggc taactacggc tacactagaa ggacagtatt	2220

	tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc	2280
5	cggcaaaaca accaccgctg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg	2340
	cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg	2400
	gaacgaaaac tcacgttaag ggatttttgt catgagatta tcaaaaagga tcttcaccta	2460
10	gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg agtaaacttg	2520
	gtctgacagt taccaatgct taatcagtg ggcacctatc tcagcgatct gtctatttcg	2580
15	ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggccttacc	2640
	atctggcccc agtgctgcaa tgataccgcg agaccacgc tcaccggctc cagattttatc	2700
	agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc	2760
20	ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag	2820
	tttgcgaaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat	2880
25	ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatggtgtg	2940
	caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt	3000
	gttatcactc atgggttatgg cagcactgca taattctctt actgtcatgc catccgtaag	3060
30	atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg	3120
	accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt	3180
35	aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct	3240
	gttgagatcc agttcgatgt aaccactcg tgcaccaaac tgatcttcag catcttttac	3300
	tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggat	3360
40	aagggcgaca cggaaatggt gaatactcat actcttctt tttcaatatt attgaagcat	3420
	ttatcagggg tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca	3480
45	aataggggtt ccgcgccacat ttccccgaaa agtgccacct gacgtctaag aaaccattat	3540
	tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc	3590
50	<210> 17 <211> 3584 <212> DNA <213> artificial sequence	
55	<220> <221> misc_feature <223> Sequenz stellt eine pflanzliche Promotor-Terminator-Expressionska ssette in Vektor pUC19 dar	
60	<400> 17 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctccc gagacggtca	60

	cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
	ttggcggggtg tcgggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
5	accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
	attcgecatt caggctgcgc aactgttggg aagggcgatc ggtgcggggc tcttcgctat	300
10	tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggg	360
	tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga	420
	gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat	480
15	gtgtgttatg tatttgattt gcgataaatt tttatatattg gtactaaatt tataacacct	540
	tttatgctaa cgtttgcaa cacttagcaa tttgcaagtt gattaattga ttctaaatta	600
20	tttttgtctt ctaaatacat atactaatca actggaaatg taaatatattg ctaatatttc	660
	tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt	720
	gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg	780
25	taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca	840
	agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt	900
30	ttaaaaaat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt	960
	ggaggatgca ataataaga aaactacaaa ttacatgca actagttatg catgtagtct	1020
	atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta	1080
35	taattttctt atagccagca gatctgccg catcgatccc gggccatggc ctgctttaat	1140
	gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg	1200
40	taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attctaata	1260
	atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa ttactgatt	1320
	gtccgtcgac gagctcggcg cgccaagctt ggcgtaatca tggcatagc tgtttcctgt	1380
45	gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa	1440
	agcctggggg gctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc	1500
50	tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag	1560
	aggcggtttg cgtattgggc gctcttcgc ttcctcgctc actgactcgc tgcgctcggt	1620
	cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga	1680
55	atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg	1740
	taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia	1800
60	aatcgcagc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt	1860
	tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct	1920

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980
 cagttcgggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040
 5 cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtta gacacgactt 2100
 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg tagggcgtgc 2160
 10 tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggtat 2220
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 2280
 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgagaaa 2340
 15 aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 2400
 aaactcacgt taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct 2460
 20 tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttgggtctga 2520
 cagttaccfaa tgcttaataca gtgaggcacc tatctcagcg atctgtctat ttcgttcac 2580
 catagtggcc tgactccccg tcgtgtagat aactacgata cgggaggggt taccatctgg 2640
 25 cccagtgct gcaatgatac cgcgagacc acgctcaccg gctccagatt tatcagcaat 2700
 aaaccagcca gccggaagg cgcgagcag aagtggctct gcaactttat ccgcctccat 2760
 30 ccagtctatt aattggttgc gggaagctag agtaagtagt tcgccagtta atagtttgcg 2820
 caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 2880
 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 2940
 35 agcggtttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc 3000
 actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt 3060
 40 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 3120
 ttgctcttgc ccggcgtaaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 3180
 gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag 3240
 45 atccagttcg atgtaacca ctcgtcacc caactgatct tcagcatctt ttactttcac 3300
 cagcgtttct gggtagcaaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 3360
 50 gacacggaaa tgttgataac tcatactctt cttttttcaa tattattgaa gcatttatca 3420
 gggttattgt ctcagagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 3480
 ggttcgcgc acatttccc gaaaagtgcc acctgacgct taagaaacca ttattatcat 3540
 55 gacattaacc tataaaaaata ggcgtatcac gaggcccttt cgtc 3584

<210> 18

<211> 4507

<212> DNA

<213> artificial sequence

<220>

<221> misc_feature

<223> Sequenz stellt eine pflanzliche Promotor-Terminator-Expressionskassette in Vektor pUC19 dar

5

<400>* 18

	tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca	60
10	cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg	120
	ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
15	accatatgcy gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
	attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggccc tcttcgctat	300
	tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
20	tttcccagtc acgacgttgt aaaacgcgcy ccagtgaatt cggcgcgccg agctcctcga	420
	gcaaattttac acattgccac taaacgtcta aacccttgta atttggtttt gttttactat	480
25	gtgtgttatg tatttgattt gcgataaatt tttatatattg gtactaaatt tataacacct	540
	tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta	600
	tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatctc	660
30	tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga ttttaattgtt	720
	gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg	780
35	taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaagggttag taatttttca	840
	agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt	900
	ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt	960
40	ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct	1020
	atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta	1080
45	taatttcttc atagccagcc caccgcggtg ggcggccgccc tgcagtctag aaggcctcct	1140
	gctttaatga gatatgcyag acgcctatga tcgcatgata tttgctttca attctgttgt	1200
	gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgccggt ttcggttcat	1260
50	tctaataaat atatcaccoc ttactatcgt atttttatga ataataattct ccgttcaatt	1320
	tactgattgt ccgtcgagca aatttacaca ttgccactaa acgtctaaac ccttgtaatt	1380
55	tgtttttggt ttactatgtg tgttatgtat ttgatttgcy ataaattttt atatttggt	1440
7	ctaaatttat aacacctttt atgctaacgt ttgccaacac ttagcaattt gcaagttgat	1500
	taattgattc taaattattt ttgtcttcta aatacatata ctaatcaact ggaaatgtaa	1560
60	atatttgcta atatttctac tataggagaa ttaaagtgag tgaatatggt accacaaggt	1620
	ttggagattt aattgttgca atgctgcatg gatggcatat acaccaaaca ttcaataatt	1680

	cttgaggata ataatggtac cacacaagat ttgaggtgca tgaacgtcac gtggacaaaa	1740
5	ggttttagtaa tttttcaaga caacaatggt accacacaca agttttgagg tgcattgcatg	1800
	gatgccctgt ggaaagtta aaaatatttt ggaaatgatt tgcattggaag ccatgtgtaa	1860
	aaccatgaca tccacttgga ggatgcaata atgaagaaaa ctacaaattt acatgcaact	1920
10	agttatgcat gtagtctata taatgaggat tttgcaatac tttcattcat acacactcac	1980
	taagttttac acgattataa tttcttcata gccagcggat ccgatatcgg gcccgctagc	2040
15	gttaaccctg ctttaatgag atatgcgaga cgcctatgat cgcattgatat ttgctttcaa	2100
	ttctgttggtg cacgttgtaa aaaacctgag catgtgtagc tcagatcctt accgccggtt	2160
	tcggttcatt ctaatgaata tatcaccctg tactatcgtt tttttatgaa taatattctc	2220
20	cgttcaattt actgattgtc cgtcgacgaa ttcgagctcg gcgcgccaag cttggcgtaa	2280
	tcattggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata	2340
25	cgagccggaa gcataaagtg taaagcctgg ggtgcctaata gattgagcta actcacatta	2400
	attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgcctgtcca gctgcattaa	2460
	tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg ggcgctcttc cgcttcctcg	2520
30	ctcactgact cgctgcgctc ggctgctcgg ctgcggcgag cggtatcagc tcaactcaaag	2580
	gcggttaatac gggtatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa	2640
35	ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc	2700
	cgccccctg acgagcatca caaaaatcga cgctcaagtc agagggtggcg aaaccgcaca	2760
	ggactataaa gataaccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg	2820
40	accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct	2880
	catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt	2940
45	gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggttaacta tcgtcttgag	3000
	tccaaccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc	3060
	agagcgagggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac	3120
50	actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga	3180
	gttggttagct cttgatccgg caaacaacc accgctggta gcgggtggtt ttttgtttgc	3240
55	aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg	3300
	gggtctgacg ctcaagtggaa cgaaaactca cgtaaggga ttttggatcat gagattatca	3360
	aaaaggatct tcacctagat ccttttaaata taaaaatgaa gttttaaatc aatctaaagt	3420
60	atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgggc acctatctca	3480
	gcgatatgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg	3540

	atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca	3600
5	ccgggtccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc	3660
	cctgcaactt tatccgctc catccagtct attaattgtt gccgggaagc tagagtaagt	3720
	agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtgggtgtca	3780
10	cgctcgtcgt ttggtatggc ttcatcagc tccgggtccc aacgatcaag gcgagttaca	3840
	tgatccccc tgttgtgcaa aaaagcgggt agctccttcg gtccctccgat cgttgtcaga	3900
15	agtaagttgg ccgcagtggt atcactcatg gttatggcag cactgcataa ttctcttact	3960
	gtcatgccat ccgtaagatg cttttctgtg actgggtgagt actcaaccaa gtcattctga	4020
	gaatagtgtg tgcggcgacc gagttgctct tgcgccggcg caatacggga taataccgcg	4080
20	ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc	4140
	tcaaggatct taccgctggt gagatccagt tcgatgtaac ccactcgtgc acccaactga	4200
25	tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat	4260
	gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt	4320
	caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt	4380
30	atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac	4440
	gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cagaggccc	4500
35	tttcgtc	4507
	<210> 19	
	<211> 17752	
	<212> DNA	
40	<213> Phaeodactylum tricornutum, Physcomitrella patens	
	<220>	
	<221> CDS	
	<222> (11543)..(12415)	
45	<223> Delta-6-Elongase	
	<220>	
	<221> CDS	
50	<222> (13313)..(14890)	
	<223> Delta-6-Desaturase	
	<220>	
	<221> CDS	
55	<222> (15791)..(17200)	
	<223> Delta-5-Desaturase	
60	<400> 19	
	gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc	60

	gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca	120
	tagtgggcggt tgacgtcggt cgagtgaacc agatcgcgca ggaggcccggt cagcaccggc	180
5	ataatcaggt cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt	240
10	atgttgggtt tcacgtctgt cctccggacc agcctccgct ggtccgattg aacgcgcgga	300
	ttcttttatca ctgataagtt ggtggacata ttatgtttat cagtataaaa gtgtcaagca	360
	tgacaaagtt gcagccgaat acagtgatcc gtgcccgcct ggacctgttg aacgaggtcg	420
	gcgtagacgg tctgacgaca cgaaaactgg cggaacgggt ggggggttcag cagccggcgc	480
15	tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg	540
	cggagaatca tacgcattcg gtgcccagag ccgacgacga ctggcgctca tttctgatcg	600
20	ggaatgccc cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg	660
	ccggcacggt accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttccctct	720
	gcgaggcggt tttttcggtt gggggacggc tcaatgcgct gatgacaatc agctacttca	780
25	ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga	840
	ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc ttcgacgaag	900
30	ccggtccgga cgcagcgttc gagcagggtc tcgcggtgat tgcgatgga ttggcgaaaa	960
	ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc	1020
	tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt	1080
35	ccaccgctc agacgccgt agcagccgc tacgggcttt ttcattgccct gccctagcgt	1140
	ccaagcctca cggccgctc cggcctctct ggcggccttc tggcgctctt ccgcttctc	1200
40	gctcactgac tcgtgcgct cggctggttc gctgcggcga gcggtatcag ctactcaaaa	1260
	ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	1320
	aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcggttt tccataggct	1380
45	ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagaggtggc gaaaccgcgac	1440
	aggactataa agataccagg cgtttcccc tggaagctcc ctctgctgct ctctgttcc	1500
50	gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt	1560
	ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt	1620
	tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttcttgg tgtatccaac	1680
55	ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca	1740
	ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg	1800
60	ctaccgcccg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga	1860
	agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa	1920

agggcgccggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980
 aaatcacggg cgctcgtggac tatgagcacg tccgcgagct ggcccgcatac aatggcgacc 2040
 5 tggggccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt 2100
 tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160
 10 gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220
 aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280
 gacttcgagg agctgggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340
 15 gacgctcacc gggctgggtg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400
 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460
 20 cctcgccgaa aacttgccc tcaactgacag atgaggggag gacgttgaca cttgaggggc 2520
 cgactcacc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580
 gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgcgagtt tcccacagat 2640
 25 gatgtggaca agcctgggga taagtgcct gcggtattga cacttgagg gcgcgactac 2700
 tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760
 30 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820
 ccgcccgttt ttggccacc gctaacctgt ctttaacct gcttttaaac caatatttat 2880
 aaacottgtt tttaaccagg gctgcgcct gtgcgcgtga ccgcgcacgc cgaagggggg 2940
 35 tgccccccct tctcgaacc tccggcccg ctaacgcggg cctcccatcc cccaggggc 3000
 tgcgcccctc ggccgcgaac ggcctcacc caaaaatggc agcgtggca gtccttgcca 3060
 40 ttgcccggat cggggcagta acgggatgg cgatcagccc gagcgcgacg cccggaagca 3120
 ttgacgtgcc gcaggtgctg gcacgacat tcagcgacca ggtgccgggc agtgagggcg 3180
 gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcagg 3240
 45 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300
 tgttcggggg tgcgataaac ccagcgaacc atttgagggt ataggtaaga ttataccgag 3360
 50 gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540
 55 ggcaaggcat aggcagcgc cttatcaata tatctataga atgggcaaag cataaaaact 3600
 tgcattgact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660
 60 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccagggt 3780

	gctgcctcag attcaggtta tgccgctcaa ttcgctgctg atatcgcttg ctgattacgt	3840
	gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca	3900
5	cgtaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga	3960
	atacgtgctc aacaaccgtc ttccggagac tgatcatagc gtaaaacagc cagcgtgtggc	4020
10	gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac	4080
	tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat	4140
	cgtgttgagg ccaacgcccc taatgcgggc tggtgcccg catccaacgc cattcatggc	4200
15	catatcaatg attttcttgt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg	4260
	ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt	4320
20	acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg	4380
	agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg	4440
	tggtttcaaa atcgggtccg tcgatactat gttatacgcc aactttgaaa acaactttga	4500
25	aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt	4560
	cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa	4620
30	taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc	4680
	gtaaaagata cggaaggaat gtctcctgct aaggatatata agctgggtggg agaaaatgaa	4740
	aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg	4800
35	gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggc cctgcacttt	4860
	gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg	4920
40	gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac	4980
	aggctctttc actccatcga catatcgatg tgtccctata cgaatagctt agacagccgc	5040
	ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg	5100
45	gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag	5160
	cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa	5220
50	gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat	5280
	gacattgcct tctgcgtccg gtgatcagg gaggatatcg gggaagaaca gtatgtcgag	5340
	ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta	5400
55	ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg	5460
	caacgacttc ttccgcatca agtgttttgg ctctcaggcc gagggccacg gcaagtattt	5520
60	gggcaagggg tcgctgggat tcgtgcaggg caagattcgg aataccaagt acgagaagga	5580
	cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa	5640

	ggcaccaggc	gggtcaaate	aggaataagg	gcacattgcc	ccggcgtgag	tcggggcaat	5700
	cccgaagga	gggtgaatga	atcggacgtt	tgaccggaag	gcatacaggc	aagaactgat	5760
5	cgacgcgggg	ttttccgccg	aggatgccga	aaccatcgca	agccgcaccg	tcattgcgtgc	5820
	gccccgcgaa	accttccagt	ccgtcggctc	gatgggtccag	caagctacgg	ccaagatcga	5880
10	gcgcgacagc	gtgcaactgg	ctccccctgc	cctgcccgcg	ccatcggccg	ccgtggagcg	5940
	ttcgcgtcgt	ctcgaacagg	aggcggcagg	tttggcgaag	tcgatgacca	tcgacacgcg	6000
	aggaactatg	acgaccaaga	agcgaaaaac	cgccggcgag	gacctggcaa	aacaggctcag	6060
15	cgaggccaag	caggccgcgt	tgctgaaaca	cacgaagcag	cagatcaagg	aaatgcagct	6120
	ttccttgctc	gatattgcgc	cgtggccgga	cacgatgcga	gcgatgccaa	acgacacggc	6180
20	ccgctctgcc	ctgttcacca	cgcgcaacaa	gaaaatcccg	cgcgaggcgc	tgcaaaacaa	6240
	ggtcattttc	cacgtcaaca	aggacgtgaa	gatcacctac	accggcgctc	agctgcgggc	6300
	cgacgatgac	gaactggtgt	ggcagcaggt	gttggagtac	gcgaagcgca	cccctatcgg	6360
25	cgagccgatc	accttcacgt	tctacgagct	ttgccaggac	ctgggctggt	cgatcaatgg	6420
	ccggtattac	acgaaggccg	aggaatgcct	gtcgcgccta	caggcgacgg	cgatgggctt	6480
30	cacgtccgac	cgcggtgggc	acctggaate	gggtgcgctg	ctgcaccgct	tcgcgcctct	6540
	ggaccgtggc	aagaaaacgt	cccgttgcca	ggctctgata	gacgaggaaa	tcgtcgtgct	6600
	gtttgctggc	gaccactaca	cgaaattcat	atgggagaag	taccgcaagc	tgctgcgcgac	6660
35	ggcccgcagg	atgttcgact	atttcagctc	gcaccgggag	ccgtacccgc	tcaagctgga	6720
	aaccttccgc	ctcatgtgcg	gatcggattc	caccgcgctg	aagaagtggc	gcgagcaggt	6780
40	cggcgaagcc	tgcaagaggt	tgcgaggcag	cggcctgggtg	gaacacgcct	gggtcaatga	6840
	tgacctgggtg	cattgcaaac	gctagggcct	tgtgggggtca	gttcgggctg	ggggttcagc	6900
	agccagcgct	ttactggcat	ttcaggaaca	agcgggcact	gctcgacgca	cttgcttcgc	6960
45	tcagtatcgc	tcgggacgca	cggcgcgctc	tacgaactgc	cgataaacag	aggattaaaa	7020
	ttgacaattg	tgattaaggc	tcagattcga	cggcttggag	cggccgacgt	gcaggatttc	7080
50	cgcgagatcc	gattgtcggc	cctgaagaaa	gctccagaga	tgttcgggtc	cgtttacgag	7140
	cacgaggaga	aaaagcccat	ggaggcgctc	gctgaacggg	tgcgagatgc	cgtggcattc	7200
	ggcgcctaca	tcgacggcga	gatcattggg	ctgtcggctc	tcaaacagga	ggacggcccc	7260
55	aaggacgctc	acaaggcgca	tctgtccggc	gttttcgtgg	agcccgaaca	gcgaggccga	7320
	ggggtcgccc	gtatgctgct	gcgggcgttg	ccggcggtt	tattgctcgt	gatgatcgtc	7380
60	cgacagattc	caacgggaat	ctgggtggatg	cgcattctca	tcctcggcgc	acttaatat	7440
	tcgctattct	ggagcttggt	gtttatttcg	gtctaccgcc	tgccgggcgg	ggtcgcggcg	7500

	acggtagggc	ctgtgcagcc	gctgatggc	gtgttcac	ctgccgctct	gctaggtagc	7560
	ccgatacgat	tgatggcgg	cctgggggct	atttgcggaa	ctgcgggcgt	ggcgctgttg	7620
5	gtgttgacac	caaacgcagc	gctagatcct	gtcggcgctc	cagcgggcct	ggcggggg	7680
	gtttccatgg	cgttcggaac	cgtgctgacc	cgcaagtggc	aacctcccgt	gcctctgctc	7740
10	acctttaccg	cctggcaact	ggcggccgga	ggacttctgc	tcgttccagt	agcttttagtg	7800
	tttgatccgc	caatcccgat	gcctacagga	accaatgttc	tcggcctggc	gtggctcggc	7860
	ctgatcggag	cgggtttaac	ctacttcctt	tggttccggg	ggatctcgcg	actcgaacct	7920
15	acagttgttt	ccttactggg	ctttctcagc	cccagatctg	gggtcgatca	gccggggatg	7980
	catcaggccg	acagtcggaa	cttcgggtcc	ccgacctgta	ccattcgggtg	agcaatggat	8040
20	aggggagttg	atatcgtcaa	cgttcacttc	taaagaaata	gcgccactca	gcttcctcag	8100
	cggcttatac	cagcgatttc	ctattatgtc	ggcatagttc	tcaagatcga	cagcctgtca	8160
	cggttaagcg	agaaatgaat	aagaaggctg	ataattcgga	tctctgcgag	ggagatgata	8220
25	tttgatcaca	ggcagcaacg	ctctgtcatc	gttacaatca	acatgctacc	ctccgcgaga	8280
	tcacccgtgt	ttcaaaccgc	gcagcttagt	tgccgttctt	ccgaatagca	tcggtaacat	8340
30	gagcaaagtc	tgccgcctta	caacggctct	cccgtgacg	ccgtcccgga	ctgatgggct	8400
	gcctgtatcg	agtgggtgatt	ttgtgccgag	ctgccggctc	gggagctggt	ggctggctgg	8460
	tggcaggata	tattgtgggtg	taaacaaatt	gacgcttaga	caacttaata	acacattgcg	8520
35	gacgttttta	atgtactggg	gtggtttttc	ttttcaccag	tgagacgggc	aacagctgat	8580
	tgcccttcac	cgcttgccc	tgagagagtt	gcagcaagcg	gtccacgctg	gtttgcccc	8640
40	gcaggcgaaa	atcctgtttg	atgggtgggtc	cgaaatcggc	aaaatccctt	ataaatcaaa	8700
	agaatagccc	gagatagggt	tgagtgttgt	tccagtttgg	aacaagagtc	cactattaaa	8760
	gaacgtggac	tccaacgtca	aagggcgaaa	aaccgtctat	cagggcgatg	gccactacg	8820
45	tgaaccatca	cccaaatcaa	gttttttggg	gtcgaggtgc	cgtaaagcac	taaatcggaa	8880
	ccctaaagg	agccccgat	ttagagcttg	acggggaaag	ccggcgaacg	tggcgagaaa	8940
50	ggaaggggag	aaagcgaaa	gagcggggcg	cattcaggct	gcgcaactgt	tgggaagggc	9000
	gatcgggtcg	ggcctcttcg	ctattacgcc	agctggcgaa	agggggatgt	gctgcaaggc	9060
	gattaagttg	ggtaacgcca	gggttttccc	agtcacgacg	ttgtaaaacg	acggccagtg	9120
55	aattaattcc	catcttgaaa	gaaatatagt	ttaaatattt	attgataaaa	taacaagtca	9180
	ggatattatag	tccaagcaaa	aacataaatt	tattgatgca	agtttaaatt	cagaaatatt	9240
60	tcaataactg	attatatcag	ctggtacatt	gccgtagatg	aaagactgag	tcgcatatta	9300
	tgtgtaatac	ataaattgat	gatatagcta	gcttagctca	tcgggggatc	cgtcgaagct	9360

agcttggggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc gccaaagctct 9480
 5 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600
 10 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaaac 9660
 agttcgggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtgggc gaatgggcag 9780
 15 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
 20 tcccttcccc cttcagtgac aacgtcgagc acagctgcgc aaggaaacgcc cgtcgtggcc 9960
 agccacgata gccgcgctgc ctgcctctgc agttcattca gggcaccgga caggctcggtc 10020
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080
 25 ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140
 cctgcgtgca atccatcttg ttcaatcaa gctcccatgg gccctcgact agagtcgaga 10200
 30 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320
 acgcgcaata atggttttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380
 35 tgagtgggtc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacgatc ttgatccct 10500
 40 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560
 cttaccagag ggcccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620
 gtctagctat cgcctatgaa gccactgca agctacgtgc tttctctttg cgcttgogtt 10680
 45 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
 actggctttc tacgtgttcc gttccttta gcagcccttg cgccctgagt gcttgccgca 10800
 50 gcgtgaagct tgcctgcctg caggctgacg gcgcgccgag ctctcgagc aaatttacac 10860
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980
 55 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100
 60 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220

	tttgagggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt	11280
	taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt	11340
5	tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat	11400
	aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga	11460
10	ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata attttcttcat	11520
	agccagccca ccgcggtgga aa atg gag gtc gtg gag aga ttc tac ggt gag	11572
	Met Glu Val Val Glu Arg Phe Tyr Gly Glu	
	1 5 10	
15	ttg gat ggg aag gtc tgc cag ggc gtg aat gca ttg ctg ggt agt ttt	11620
	Leu Asp Gly Lys Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe	
	15 20 25	
20	ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctg gtt	11668
	Gly Val Glu Leu Thr Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val	
	30 35 40	
	gac agt ccc aca ccc atc gtc ctg ggt gtt tct gta tac ttg act att	11716
25	Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile	
	45 50 55	
	gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc	11764
	Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg	
30	60 65 70	
	gcc tcg gag cca ttt ttg ctg caa gct ttg gtg ctt gtg cac aac ctg	11812
	Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu	
	75 80 85 90	
35	ttc tgt ttt gcg ctg agt ctg tat atg tgc gtg ggc atc gct tat cag	11860
	Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln	
	95 100 105	
40	gct att acc tgg cgg tac tct ctg tgg ggc aat gca tac aat cct aaa	11908
	Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys	
	110 115 120	
	cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac	11956
45	His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr	
	125 130 135	
	gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg	12004
	Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg	
50	140 145 150	
	caa ata agc ttc ctg cac gtt tat cat cat tct tca att tcc ctg att	12052
	Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Ser Leu Ile	
	155 160 165 170	
55	tgg tgg gct att gct cat cac gct cct ggc ggt gaa gca tat tgg tct	12100
	Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser	
	175 180 185	
60	gcg gct ctg aac tca gga gtg cat gtt ctg atg tat gcg tat tac ttc	12148
	Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe	
	190 195 200	

	ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt	12196
	Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu	
	205 210 215	
5	ttt tgg ggc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg	12244
	Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu	
	220 225 230	
10	aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca	12292
	Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro	
	235 240 245 250	
15	caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt	12340
	Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe	
	255 260 265	
20	ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga	12388
	Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly	
	270 275 280	
	aag caa aag gga gct aaa act gag tga tctagaaggc ctctgcttt	12435
	Lys Gln Lys Gly Ala Lys Thr Glu	
	285 290	
25	aatgagatat gcgagacgcc tatgatcgca tgatatattgc tttcaattct gttgtgcacg	12495
	ttgtaaaaaa cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa	12555
30	tgaatatatc acccgttact atcgatatttt tatgaataat attctccgtt caatttactg	12615
	attgtccgtc gagcaaatTT acacattgcc actaaacgtc taaacccttg taatttgttt	12675
	ttgttttact atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tgggtactaaa	12735
35	tttataacac cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt	12795
	gattctaaat tattttttgtc ttctaaatac atatactaata caactggaaa tgtaaataatt	12855
40	tgctaataatt tctactatag gagaattaaa gtgagtgaat atgggtaccac aagggtttgga	12915
	gatttaattg ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga	12975
	ggataataat ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaagggttt	13035
45	agtaattttt caagacaaca atgttaccac acacaagttt tgagggtgcat gcatggatgc	13095
	cctgtggaaa gtttaaaaat attttggaat tgatttgcat ggaagccatg tgtaaaacca	13155
50	tgacatccac ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta	13215
	tgcatgtagt ctatataatg aggattttgc aatactttca ttcatacaca ctactaagt	13275
	tttacacgat tataatttct tcatagccag cggatcc atg gta ttc gcg ggc ggt	13330
55		Met Val Phe Ala Gly Gly
		295
60	gga ctt cag cag ggc tct ctc gaa gaa aac atc gac gtc gag cac att	13378
	Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn Ile Asp Val Glu His Ile	
	300 305 310	
	gcc agt atg tct ctc ttc agc gac ttc ttc agt tat gtg tct tca act	13426
	Ala Ser Met Ser Leu Phe Ser Asp Phe Phe Ser Tyr Val Ser Ser Thr	

	315	320	325	
5	gtt ggt tcg tgg agc gta cac agt ata caa cct ttg aag cgc ctg acg Val Gly Ser Trp Ser Val His Ser Ile Gln Pro Leu Lys Arg Leu Thr 330 335 340			13474
10	agt aag aag cgt gtt tcg gaa agc gct gcc gtg caa tgt ata tca gct Ser Lys Lys Arg Val Ser Glu Ser Ala Ala Val Gln Cys Ile Ser Ala 345 350 355 360			13522
	gaa gtt. cag aga aat tcg agt acc cag gga act gcg gag gca ctc gca Glu Val Gln Arg Asn Ser Ser Thr Gln Gly Thr Ala Glu Ala Leu Ala 365 370 375			13570
15	gaa tca gtc gtg aag ccc acg aga cga agg tca tct cag tgg aag aag Glu Ser Val Val Lys Pro Thr Arg Arg Arg Ser Ser Gln Trp Lys Lys 380 385 390			13618
20	tcg aca cac ccc cta tca gaa gta gca gta cac aac aag cca agc gat Ser Thr His Pro Leu Ser Glu Val Ala Val His Asn Lys Pro Ser Asp 395 400 405			13666
25	tgc tgg att gtt gta aaa aac aag gtg tat gat gtt tcc aat ttt gcg Cys Trp Ile Val Val Lys Asn Lys Val Tyr Asp Val Ser Asn Phe Ala 410 415 420			13714
30	gac gag cat ccc gga gga tca gtt att agt act tat ttt gga cga gac Asp Glu His Pro Gly Gly Ser Val Ile Ser Thr Tyr Phe Gly Arg Asp 425 430 435 440			13762
	ggc aca gat gtt ttc tct agt ttt cat gca gct tct aca tgg aaa att Gly Thr Asp Val Phe Ser Ser Phe His Ala Ala Ser Thr Trp Lys Ile 445 450 455			13810
35	ctt caa gac ttt tac att ggt gac gtg gag agg gtg gag ccg act cca Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu Arg Val Glu Pro Thr Pro 460 465 470			13858
40	gag ctg ctg aaa gat ttc cga gaa atg aga gct ctt ttc ctg agg gag Glu Leu Leu Lys Asp Phe Arg Glu Met Arg Ala Leu Phe Leu Arg Glu 475 480 485			13906
45	caa ctt ttc aaa agt tcg aaa ttg tac tat gtt atg aag ctg ctc acg Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr Val Met Lys Leu Leu Thr 490 495 500			13954
50	aat gtt gct att ttt gct gcg agc att gca ata ata tgt tgg agc aag Asn Val Ala Ile Phe Ala Ala Ser Ile Ala Ile Cys Trp Ser Lys 505 510 515 520			14002
	act att tca gcg gtt ttg gct tca gct tgt atg atg gct ctg tgt ttc Thr Ile Ser Ala Val Leu Ala Ser Ala Cys Met Met Ala Leu Cys Phe 525 530 535			14050
55	caa cag tgc gga tgg cta tcc cat gat ttt ctc cac aat cag gtg ttt Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His Asn Gln Val Phe 540 545 550			14098
60	gag aca cgc tgg ctt aat gaa gtt gtc ggg tat gtg atc ggc aac gcc Glu Thr Arg Trp Leu Asn Glu Val Val Gly Tyr Val Ile Gly Asn Ala 555 560 565			14146

	gtt ctg ggg ttt agt aca ggg tgg tgg aag gag aag cat aac ctt cat	14194
	Val Leu Gly Phe Ser Thr Gly Trp Trp Lys Glu Lys His Asn Leu His	
	570 575 580	
5	cat gct gct cca aat gaa tgc gat cag act tac caa cca att gat gaa	14242
	His Ala Ala Pro Asn Glu Cys Asp Gln Thr Tyr Gln Pro Ile Asp Glu	
	585 590 595 600	
10	gat att gat act ctc ccc ctc att gcc tgg agc aag gac ata ctg gcc	14290
	Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp Ser Lys Asp Ile Leu Ala	
	605 610 615	
15	aca gtt gag aat aag aca ttc ttg cga atc ctc caa tac cag cat ctg	14338
	Thr Val Glu Asn Lys Thr Phe Leu Arg Ile Leu Gln Tyr Gln His Leu	
	620 625 630	
20	ttc ttc atg ggt ctg tta ttt ttc gcc cgt ggt agt tgg ctc ttt tgg	14386
	Phe Phe Met Gly Leu Leu Phe Phe Ala Arg Gly Ser Trp Leu Phe Trp	
	635 640 645	
	agc tgg aga tat acc tct aca gca gtg ctc tca cct gtc gac agg ttg	14434
	Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu Ser Pro Val Asp Arg Leu	
	650 655 660	
25	ttg gag aag gga act gtt ctg ttt cac tac ttt tgg ttc gtc ggg aca	14482
	Leu Glu Lys Gly Thr Val Leu Phe His Tyr Phe Trp Phe Val Gly Thr	
	665 670 675 680	
30	gcg tgc tat ctt ctc cct ggt tgg aag cca tta gta tgg atg gcg gtg	14530
	Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro Leu Val Trp Met Ala Val	
	685 690 695	
35	act gag ctc atg tcc ggc atg ctg ctg ggc ttt gta ttt gta ctt agc	14578
	Thr Glu Leu Met Ser Gly Met Leu Leu Gly Phe Val Phe Val Leu Ser	
	700 705 710	
40	cac aat ggg atg gag gtt tat aat tgc tct aaa gaa ttc gtg agt gca	14626
	His Asn Gly Met Glu Val Tyr Asn Ser Ser Lys Glu Phe Val Ser Ala	
	715 720 725	
	cag atc gta tcc aca cgg gat atc aaa gga aac ata ttc aac gac tgg	14674
	Gln Ile Val Ser Thr Arg Asp Ile Lys Gly Asn Ile Phe Asn Asp Trp	
	730 735 740	
45	ttc act ggt ggc ctt aac agg caa ata gag cat cat ctt ttc cca aca	14722
	Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro Thr	
	745 750 755 760	
50	atg ccc agg cat aat tta aac aaa ata gca cct aga gtg gag gtg ttc	14770
	Met Pro Arg His Asn Leu Asn Lys Ile Ala Pro Arg Val Glu Val Phe	
	765 770 775	
55	tgt aag aaa cac ggt ctg gtg tac gaa gac gta tct att gct acc ggc	14818
	Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Ile Ala Thr Gly	
	780 785 790	
60	act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca	14866
	Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala	
	795 800 805	
	gag cag cat gct acc acc agt taa gctagcgtta accctgctttt aatgagatat	14920
	Glu Gln His Ala Thr Thr Ser	

810	815	
	gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa	14980
5	cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc	15040
	accogttact atcgtatttt tatgaataat attctccgtt caatttactg attgtccgtc	15100
10	gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt ttgttttact	15160
	atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tgggtactaaa tttataacac	15220
	cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat	15280
15	tatttttgtc ttctaaatac atatactaata caactggaaa tgtaaataatt tgctaataatt	15340
	tctactatag gagaattaaa gtgagtgaat atgggtaccac aagggtttgga gatttaattg	15400
20	ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga ggataataat	15460
	ggtaccaacac aagatttgag gtgcatgaac gtcacgtgga caaaagggtt agtaattttt	15520
	caagaqaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa	15580
25	gtttaaaaat attttggaat tgatttgcat ggaagccatg tgtaaaacca tgacatccac	15640
	ttggaggatg caataatgaa gaaaactaca aatttactg caactagtta tgcattgtatg	15700
30	ctatataatg aggattttgc aatactttca ttcatacaca ctactaagt tttacacgat	15760
	tataatttct tcatagccag cagatctaaa atg gct ccg gat gcg gat aag ctt	15814
	Met Ala Pro Asp Ala Asp Lys Leu	820
35	cga caa cgc cag acg act gcg gta gcg aag cac aat gct gct acc ata	15862
	Arg Gln Arg Gln Thr Thr Ala Val Ala Lys His Asn Ala Ala Thr Ile	825 830 835
40	tcg acg cag gaa cgc ctt tgc agt ctg tct tcg ctc aaa ggc gaa gaa	15910
	Ser Thr Gln Glu Arg Leu Cys Ser Leu Ser Ser Leu Lys Gly Glu Glu	840 845 850 855
45	gtc tgc atc gac gga atc atc tat gac ctc caa tca ttc gat cat ccc	15958
	Val Cys Ile Asp Gly Ile Ile Tyr Asp Leu Gln Ser Phe Asp His Pro	860 865 870
50	ggg ggt gaa acg atc aaa atg ttt ggt ggc aac gat gtc act gta cag	16006
	Gly Gly Glu Thr Ile Lys Met Phe Gly Gly Asn Asp Val Thr Val Gln	875 880 885
	tac aag atg att cac ccg tac cat acc gag aag cat ttg gaa aag atg	16054
	Tyr Lys Met Ile His Pro Tyr His Thr Glu Lys His Leu Glu Lys Met	890 895 900
55	aag cgt gtc ggc aag gtg acg gat ttc gtc tgc gag tac aag ttc gat	16102
	Lys Arg Val Gly Lys Val Thr Asp Phe Val Cys Glu Tyr Lys Phe Asp	905 910 915
60	acc gaa ttt gaa cgc gaa atc aaa cga gaa gtc ttc aag att gtg cga	16150
	Thr Glu Phe Glu Arg Glu Ile Lys Arg Glu Val Phe Lys Ile Val Arg	920 925 930 935

	cga ggc aag gat ttc ggt act ttg gga tgg ttc ttc cgt gcg ttt tgc	16198
	Arg Gly Lys Asp Phe Gly Thr Leu Gly Trp Phe Phe Arg Ala Phe Cys	
	940 945 950	
5	tac att gcc att ttc ttc tac ctg cag tac cat tgg gtc acc acg gga	16246
	Tyr Ile Ala Ile Phe Phe Tyr Leu Gln Tyr His Trp Val Thr Thr Gly	
	955 960 965	
10	acc tct tgg ctg ctg gcc gtg gcc tac gga atc tcc caa gcg atg att	16294
	Thr Ser Trp Leu Leu Ala Val Ala Tyr Gly Ile Ser Gln Ala Met Ile	
	970 975 980	
15	ggc atg aat gtc cag cac gat gcc aac cac ggg gcc acc tcc aag cgt	16342
	Gly Met Asn Val Gln His Asp Ala Asn His Gly Ala Thr Ser Lys Arg	
	985 990 995	
20	ccd tgg gtc aac gac atg cta ggc ctc ggt gcg gat ttt att ggt	16387
	Pro Trp Val Asn Asp Met Leu Gly Leu Gly Ala Asp Phe Ile Gly	
	1000 1005 1010	
	ggg tcc aag tgg ctc tgg cag gaa caa cac tgg acc cac cac gct	16432
	Gly Ser Lys Trp Leu Trp Gln Glu Gln His Trp Thr His His Ala	
	1015 1020 1025	
25	tac acc aat cac gcc gag atg gat ccc gat agc ttt ggt gcc gaa	16477
	Tyr Thr Asn His Ala Glu Met Asp Pro Asp Ser Phe Gly Ala Glu	
	1030 1035 1040	
30	cca atg ctc cta ttc aac gac tat ccc ttg gat cat ccc gct cgt	16522
	Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp His Pro Ala Arg	
	1045 1050 1055	
35	acc tgg cta cat cgc ttt caa gca ttc ttt tac atg ccc gtc ttg	16567
	Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met Pro Val Leu	
	1060 1065 1070	
40	gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att ctt gac	16612
	Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile Leu Asp	
	1075 1080 1085	
	ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac aac	16657
	Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp Asn	
	1090 1095 1100	
45	gct ttc att cac tcg cga cgc aag tat gcg gtt ttc tgg cgg gct	16702
	Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala	
	1105 1110 1115	
50	gtg tac att gcg gtg aac gtg att gct ccg ttt tac aca aac tcc	16747
	Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser	
	1120 1125 1130	
55	ggc ctc gaa tgg tcc tgg cgt gtc ttt gga aac atc atg ctc atg	16792
	Gly Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met	
	1135 1140 1145	
60	ggg gtg gcg gaa tcg ctc gcg ctg gcg gtc ctg ttt tcg ttg tcg	16837
	Gly Val Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser	
	1150 1155 1160	
	cac aat ttc gaa tcc gcg gat cgc gat ccg acc gcc cca ctg aaa	16882
	His Asn Phe Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys	

	1165		1170		1175		
5	aag Lys 1180	acg gga gaa cca gtc Thr Gly Glu Pro Val 1185	gac tgg ttc aag aca Asp Trp Phe Lys Thr 1190	cag gtc gaa act Gln Val Glu Thr	16927		
10	tcc Ser 1195	tgc act tac ggt gga Cys Thr Tyr Gly Gly 1200	ttc ctt tcc ggt tgc Phe Leu Ser Gly Cys 1205	ttc acg gga ggt Phe Thr Gly Gly	16972		
	ctc Leu 1210	aac ttt cag gtt gaa Asn Phe Gln Val Glu 1215	cac cac ttg ttc cca His His Leu Phe Pro 1220	cgc atg agc agc Arg Met Ser Ser	17017		
15	gct Ala 1225	tgg tat ccc tac att Trp Tyr Pro Tyr Ile 1230	gcc ccc aag gtc cgc Ala Pro Lys Val Arg 1235	gaa att tgc gcc Glu Ile Cys Ala	17062		
20	aaa Lys 1240	cac ggc gtc cac tac His Gly Val His Tyr 1245	gcc tac tac ccg tgg Ala Tyr Tyr Pro Trp 1250	atc cac caa aac Ile His Gln Asn	17107		
25	ttt Phe 1255	ctc tcc acc gtc cgc Leu Ser Thr Val Arg 1260	tac atg cac gcg gcc Tyr Met His Ala Ala 1265	ggg acc ggt gcc Gly Thr Gly Ala	17152		
30	aac Asn 1270	tgg cgc cag atg gcc Trp Arg Gln Met Ala 1275	aga gaa aat ccc ttg Arg Glu Asn Pro Leu 1280	acc gga cgg gcg Thr Gly Arg Ala	17197		
	taa agatctgccg gcatcgatcc	cgggccatgg cctgctttaa	tgagatatgc		17250		
35	gagacgccta tgatcgcatg	atatttgctt tcaattctgt	tgtgcacggt gtaaaaaacc		17310		
	tgagcatgtg tagctcagat	ccttaccgcc gggttcggtt	cattctaattg aatatatcac		17370		
	ccgttactat cgtatTTTTTA	tgaataatat tctccgttca	atttactgat tgtccgtcga		17430		
40	cgagctcggc ggcctctag	aggatcgatg aattcagatc	ggctgagtgg ctccctcaac		17490		
	gttgcggttc tgtcagttcc	aaacgtaaaa cggttggtcc	cgcgtcacgc gcgggggtca		17550		
	taacgtgact cccttaattc	tccgctcatg atcagattgt	cgtttccgc cttcagttta		17610		
45	aactatcagt gtttgacagg	atatattggc gggtaaacct	aagagaaaag agcgtttatt		17670		
	agaataatcg gatattttaa	agggcgtgaa aaggtttata	cttcgtccat ttgtatgtgc		17730		
50	atgccaaacca cagggttccc	ca			17752		
	<210> 20						
	<211> 290						
	<212> PRT						
55	<213> Phaeodactylum tricornutum, Physcomitrella patens						
	<400> 20						
60	Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser						
	1	5	10	15			

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
 20 25 30

5 Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
 35 40 45

10 Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
 50 55 60

15 Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
 65 70 75 80

20 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95

25 Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr
 100 105 110

30 Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
 115 120 125

35 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
 130 135 140

40 Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
 145 150 155 160

45 Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
 165 170 175

50 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
 180 185 190

55 Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
 195 200 205

60 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
 210 215 220

65 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
 225 230 235 240

70 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
 245 250 255

75 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
 260 265 270

5 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
 275 280 285
 Thr Glu
 290
 10 <210> 21
 <211> 525
 <212> PRT
 <213> Phaeodactylum tricornutum, Physcomitrella patens
 15 <400> 21
 20 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn
 1 5 10 15
 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe
 20 25 30
 25 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln
 35 40 45
 30 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala
 50 55 60
 35 Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly
 65 70 75 80
 40 Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg
 85 90 95
 Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val
 100 105 110
 45 His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr
 115 120 125
 50 Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser
 130 135 140
 55 Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala
 145 150 155 160
 60 Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu
 165 170 175
 Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg

	180	185	190
5	Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr 195 200 205		
10	Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala 210 215 220		
15	Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys 225 230 235 240		
20	Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe 245 250 255		
25	Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly 260 265 270		
30	Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys 275 280 285		
35	Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr 290 295 300		
40	Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp 305 310 315 320		
45	Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile 325 330 335		
50	Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg 340 345 350		
55	Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu 355 360 365		
60	Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr 370 375 380		
	Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro 385 390 395 400		
	Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly 405 410 415		
	Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser 420 425 430		

Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly
 435 440 445
 5
 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu
 450 455 460
 10 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala
 465 470 475 480
 15 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp
 485 490 495
 20 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu
 500 505 510
 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser
 515 520 525
 25
 <210> 22
 <211> 469
 <212> PRT
 <213> *Phaeodactylum tricornutum*, *Physcomitrella patens*
 30
 <400> 22
 Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val
 1 5 10 15
 35 Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser
 20 25 30
 40 Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr
 35 40 45
 45 Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe
 50 55 60
 50 Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His
 65 70 75 80
 Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp
 85 90 95
 55 Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys
 100 105 110
 60 Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu
 115 120 125

5 Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu
 130 135 140
 10 Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala
 145 150 155 160
 15 Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala
 165 170 175
 20 Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly
 180 185 190
 25 Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln
 195 200 205
 30 His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp
 210 215 220
 35 Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp
 225 230 235 240
 40 His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met
 245 250 255
 45 Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile
 260 265 270
 50 Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp
 275 280 285
 55 Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala
 290 295 300
 60 Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly
 305 310 315 320
 65 Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val
 325 330 335
 70 Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe
 340 345 350
 75 Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu
 355 360 365

Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly
 370 375 380

5 Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu
 385 390 395 400

10 His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala
 405 410 415

15 Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr
 420 425 430

20 Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His
 435 440 445

Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro
 450 455 460

25 Leu Thr Gly Arg Ala
 465

30 <210> 23
 <211> 26
 <212> DNA
 <213> artificial sequence

35 <400> 23
 gaattcggcg cgccgagctc ctcgag 26

40 <210> 24
 <211> 265
 <212> DNA
 <213> artificial sequence

45 <400> 24
 ccaccgcggt gggcggccgc ctgcagtcta gaaggcctcc tgctttaatg agatatgcga 60
 gacgcctatg atcgcatgat atttgctttc aattctgttg tgcacgttgt aaaaaacctg 120
 agcatgtgta gctcagatcc ttaccgccgg ttctggttca ttctaataa tatatcaccc 180

50 gttactatcg tatttttatg aataaatattc tccgttcaat ttactgattg tccgtcgacg 240
 aattcgagct cggcgcgcca agctt 265

55 <210> 25
 <211> 257
 <212> DNA
 <213> artificial sequence

60 <400> 25
 ggatccgata tcgggcccgc tagcgtaaac cctgctttaa tgagatatgc gagacgccta 60

	tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc tgagcatgtg	120
	tagctcagat ccttaccgcc gggttcgggtt cattctaataa aatatatcac ccgttactat	180
5	cgtatttttta tgaataatat tctccgttca atttactgat tgtccgtcga cgaattcgag	240
	ctcggcgccg caagctt	257
10	<210> 26 <211> 5410 <212> DNA <213> artificial sequence	
15	<400> 26 ttttggaaat gatttgcatg gaagccatgt gtaaaaccat gacatccact tggaggatgc	60
	aataatgaag aaaactacaa atttacatgc aactagttat gcatgtagtc tatataatga	120
20	ggattttgca atactttcat tcatacacac tcaactaagtt ttacacgatt ataatttctt	180
	catagccagc ggatccgata tcggggccgc tagcgttaac cctgctttta tgagatatgc	240
	gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc	300
25	tgagcatgtg tagctcagat ccttaccgcc gggttcgggtt cattctaataa aatatatcac	360
	ccgttactat cgtatttttta tgaataatat tctccgttca atttactgat tgtccgtcga	420
30	gcaaattttac acattgccac taaacgtcta aacccttgta atttggtttt gttttactat	480
	gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct	540
	tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta	600
35	tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc	660
	tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgagga ttttaattgtt	720
40	gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg	780
	taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca	840
	agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgccc tgtggaaagt	900
45	ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt	960
	ggaggatgca ataataaga aaactacaaa ttacatgca actagttatg catgtagtct	1020
50	atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta	1080
	taattttctt atagccagca gatctgccg catcgatccc gggccatggc ctgctttaat	1140
	gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg	1200
55	taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcggttc attctaataa	1260
	atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa tttactgatt	1320
60	gtccgtcgac gagctcggcg cgccaagctt ggcgtaatca tggatcatagc tgtttcctgt	1380
	gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa	1440

agcctgggggt gcctaagttag tgagctaact cacattaatt gcgttgcgct cactgcccgc 1500
 5 tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaaac gcgcggggag 1560
 aggcgggtttg cgtattgggc gctcttccgc ttctctgctc actgactcgc tgcgctcggt 1620
 cgttcgggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 1680
 10 atcaggggat aacgcaggaa agaactatgtg agcaaaaggc cagcaaaagg ccaggaaccg 1740
 taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia 1800
 15 aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt 1860
 tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 1920
 gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980
 20 cagttcgggtg taggtcggtc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040
 cgacegetgc gccttatccg gtaactatcg tcttgagtc aaccgggtaa gacacgactt 2100
 25 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc 2160
 tacagagttc ttgaagtggg gccctaacta cggctacact agaaggacag tatttggtat 2220
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 2280
 30 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgagaaa 2340
 aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaaacga 2400
 35 aaactcacgt taagggattt tgggtcatgag attatcaaaa aggatcttca cctagatcct 2460
 tttaaattaa aaatgaagtt ttaaataat cttaaagtata tatgagtaaa cttgggtctga 2520
 cagttaccaa tgcttaata gtagggcacc tatctcagcg atctgtctat ttcgttcac 2580
 40 catagttgcc tgactccccg tcgtgtagat aactacgata cgggaggggt taccatctgg 2640
 cccagtgct gcaatgatac cgcgagacc acgctcaccg gctccagatt tatcagcaat 2700
 45 aaaccagcca gccggaagg cgcgagcag aagtggtcct gcaactttat ccgcctccat 2760
 ccagtctatt aattggtgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg 2820
 caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 2880
 50 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 2940
 agcggttagc tccttcggtc ctccgatcgt tgcagaagt aagttggccg cagtgttatc 3000
 actcatgggt atggcagcac tgcataatc tcttactgtc atgccatccg taagatgctt 3060
 55 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 3120
 ttgctcttgc ccggcgtaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 3180
 60 gctcatcatt ggaaaacgt cttcggggcg aaaactctca aggatcttac cgctgttgag 3240
 atccagttcg atgtaacca ctcgtcacc caactgatct tcagcatctt ttactttcac 3300

	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	3360
5	gacacggaaa	tgttgaatac	tcatactctt	cctttttcaa	tattattgaa	gcatttatca	3420
	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	3480
	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgtc	taagaaacca	ttattatcat	3540
10	gacattaacc	tataaaaata	ggcgtatcac	gaggcccttt	cgtctcgcgc	gtttcgggtga	3600
	tgacgggtgaa	aacctctgac	acatgcagct	cccggagacg	gtcacagctt	gtctgtaagc	3660
15	ggatgccggg	agcagacaag	cccgtcaggg	cgcgtcagcg	gggtgtggcg	gggtgtcgggg	3720
	ctggcttaac	tatgcggcat	cagagcagat	tgtactgaga	gtgcaccata	tgccgtgtga	3780
	aataccgcac	agatgcgtaa	ggagaaaata	ccgcatcagg	cgccattcgc	cattcaggct	3840
20	gcgcaactgt	tgggaagggc	gatcggtgcg	ggcctcttcg	ctattacgcc	agctggcgaa	3900
	agggggatgt	gctgcaaggc	gattaagttg	ggtaacgcc	gggttttccc	agtcacgacg	3960
25	ttgtaaaacg	acggccagtg	aattcggcgc	gdcgagctcc	tcgagcaa	ttacacattg	4020
	ccactaaacg	tctaaaccct	tgtaatttgt	ttttgtttta	ctatgtgtgt	tatgtatttg	4080
	atgtgcgata	aattttttata	tttggtacta	aattttataac	accttttatg	ctaactgttg	4140
30	ccaacactta	gcaatttgca	agttgattaa	ttgattctaa	attatttttg	tcttctaaat	4200
	acataacta	atcaactgga	aatgtaaata	tttgctaata	tttctactat	aggagaatta	4260
35	aagtgagtga	atatggtacc	acaaggtttg	gagattta	tggtgcaatg	ctgcatggat	4320
	ggcatataca	ccaaacattc	aataattctt	gaggataata	atggtaccac	acaagatttg	4380
	aggtgcatga	acgtcacgtg	gacaaaaggt	ttagtaattt	ttcaagacaa	caatgttacc	4440
40	acacacaagt	tttgaggtgc	atgcatggat	gccctgtgga	aagtttaaaa	atattttgga	4500
	aatgatttgc	atggaagcca	tgtgtaaaac	catgacatcc	acttgaggga	tgcaataatg	4560
45	aagaaaacta	caaatttaca	tgcaactagt	tatgcatgta	gtctatataa	tgaggatttt	4620
	gcaatacttt	cattcataca	cactcactaa	gtttttacacg	attataattt	cttcatagcc	4680
	agcccaccgc	ggtgggcggc	cgcctgcagt	ctagaaggcc	tcctgcttta	atgagatatg	4740
50	cgagacgcct	atgatcgcat	gatatttgct	ttcaattctg	ttgtgcacgt	tgtaaaaaac	4800
	ctgagcatgt	gtagctcaga	tccttaccgc	cggtttcggg	tcattcta	gaatatatca	4860
55	cccgttacta	tcgtattttt	atgaataata	ttctccgttc	aatttactga	ttgtccgtcg	4920
	agcaaattta	cacattgcca	ctaaacgtct	aaacccttgt	aatttgtttt	tgttttacta	4980
	tgtgtgttat	gtatttgatt	tgcgataaat	ttttatattt	ggtactaaat	ttataacacc	5040
60	ttttatgcta	acgtttgcca	acacttagca	atttgcaagt	tgattaattg	attctaaatt	5100
	atttttgtct	tctaaataca	tatactaata	aactggaaat	gtaaatattt	gctaataattt	5160

	ctactatagg agaattaaag tgagtgaata tggtaggaccaca aggtttggag atttaattgt	5220
5	tgcaatgctg catggatggc atatacacca aacattcaat aattcttgag gataataatg	5280
	gtaccacaca agatttgagg tgcatagaacg tcacgtggac aaaagggttta gtaatttttc	5340
	aagacaacaa tgttaggaccaca cacaagtttt gaggtgcatg catggatgcc ctgtggaaag	5400
10	tttaaaaata	5410
	<210> 27	
	<211> 12093	
15	<212> DNA	
	<213> artificial sequence	
	<400> 27	
20	gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc	60
	gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca	120
	tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc	180
25	ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt	240
	atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga	300
30	ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaaa gtgtcaagca	360
	tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg	420
	gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc	480
35	tttactggca cttcaggaac aagcgggagc tgctcgacgc actggccgaa gccatgctgg	540
	cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg	600
40	ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg	660
	ccggcacgcg accgggagca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct	720
	gcgaggcggg tttttcggcc ggggacgccc tcaatgcgct gatgacaatc agctacttca	780
45	ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca	840
	ccgttgaaca ggctccgctc tcgccgctgt tgcgggcccgc gatagacgcc ttcgacgaag	900
50	ccggtccgga cgcagcgttc gagcaggagc tcgcggtgat tgcgatgga ttggcgaaaa	960
	ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc	1020
	tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt	1080
55	ccaccgcgtc agacgcccgt agcagcccgc tacgggcttt ttcattgccct gccctagcgt	1140
	ccaagcctca cggccgagct cggcctctct ggcggccttc tggcgctctt ccgcttctc	1200
60	gctcactgac tcgctgcgct cggctgctcg gctgcggcga gcggtatcag ctactcaaaa	1260
	ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	1320

	aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	1380
	ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac	1440
5	aggactataa agataccagg cgtttcccc tggaagetcc ctctgctgct ctctgttcc	1500
	gacctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt	1560
10	ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt	1620
	tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttccttgg tgtatccaac	1680
	ggcgtcagcc gggcaggata ggtgaagtag gccaccgc gagcgggtgt tccttcttca	1740
15	ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg	1800
	ctaaccgccc cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga	1860
20	agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa	1920
	agggcgccgc gggcggcatg agcctgtcgg cctacctgct ggccgctcggc cagggtaca	1980
	aaatcapggg cgctgtggac tatgagcacg tccgcgagct ggcccgcac aatggcgacc	2040
25	tgggcccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt	2100
	tcggtgatgc cagatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg	2160
30	gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa	2220
	aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc	2280
	gacttcgagg agctgggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc	2340
35	gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa	2400
	cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata	2460
40	cctcgcgga aacttgcccc tcaactgacag atgaggggag gacgttgaca cttgaggggc	2520
	cgactcacc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg	2580
	gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgcgagtt tcccacagat	2640
45	gatgtggaca agcctgggga taagtgcct gcggtattga cacttgaggg gcgcgactac	2700
	tgacagatga ggggcgcgat ccttgacact tgaggggag agtgctgaca gatgaggggc	2760
50	gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggtt	2820
	ccgcccgttt ttcgccacc gctaacctgt cttttaacct gcttttaaac caatatttat	2880
	aaaccttggt tttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg	2940
55	tgccccccct tctcgaacct tcccggccc ctaacgcggg cctcccatcc cccaggggc	3000
	tgcgcccctc ggcgcgaac ggctcacc caaaaatggc agcgtggca gtccttgcca	3060
60	ttgcccggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca	3120
	ttgacgtgcc gcagggtgct gcatcgacat tcagcgacca ggtgccgggc agtgagggcg	3180

gcggcctggg tggcggcctg cccttcactt cggccgctcg ggcattcacg gacttcatgg 3240
 cggggccggc aattttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300
 5 tgttcggggg tgcgataaac ccagcgaacc atttgagggtg ataggtaaga ttataccgag 3360
 gtatgaaaac gagaattgga cttttacaga attactctat gaagcgccat atttaaaaag 3420
 10 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600
 15 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720
 20 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccagggt 3780
 gctgcctcag attcagggtta tgccgctcaa ttccgtcgtt atatcgcttg ctgattacgt 3840
 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900
 25 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020
 30 gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac 4080
 tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140
 cgtgttgagg ccaacgcca taatgcgggc tggtgccgg catccaacgc cattcatggc 4200
 35 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320
 40 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg 4440
 tggtttcaaa atcgggctccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500
 45 aaaagctggt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620
 50 taaatggcta aatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaagata cggaaggaat gtctcctgct aagggtatata agctgggtggg agaaaatgaa 4740
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800
 55 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920
 60 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggtcttttc actccatcga catatcggtat tgtccctata cgaatagctt agacagccgc 5040

ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag 5160
 5 cccgaagagg aacttgtctt tcccacggc gacctgggag acagcaacat ctttgtgaaa 5220
 gatggcaaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280
 10 gacattgctt tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340
 ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460
 15 caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggccacg gcaagtattt 5520
 gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580
 20 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
 ggcaccaggc gggtaaatc aggaataagg gcacattgcc ccggcgtag tcggggcaat 5700
 cccgcaagga gggtagaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760
 25 cgacgcgggg tttccgcgcg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820
 gccccgcgaa accttccagt ccgtcggctc gatgggccag caagctacgg ccaagatcga 5880
 30 gcgcgacacg gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940
 ttcgcgtcgt ctcgaaacagg agggcgagc tttggcgaag tcgatgacca tcgacacgcg 6000
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacagggtcag 6060
 35 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120
 ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc 6180
 40 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgctc agctgcgggc 6300
 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360
 45 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480
 50 cacgtccgac cgcgttgggc acctggaatc ggtgtcgtc ctgcaccgct tccgcgtcct 6540
 ggaccgtggc aagaaaacgt ccggttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600
 gtttgcgtggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660
 55 gggccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720
 aaacttccgc ctcatgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780
 60 cggcgaagcc tgcaagagt tgcgaggcag cggcctggtg gaacacgcct gggtaaatga 6840
 tgacctggtg cattgcaaac gctagggcct tgtgggtca gttccggctg ggggttcagc 6900

	agccagcgct ttactggcat ttcaggaaca agcggggcact gctcgacgca cttgcttcgc	6960
	tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa	7020
5	ttgacaattg tgattaaggc tcagattcga cggcttgagg cggccgacgt gcaggatttc	7080
	cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcggggtc cgtttacgag	7140
10	cacgaggaga aaaagcccat ggaggcgctc gctgaacggt tgcgagatgc cgtggcattc	7200
	ggcgccctaca tcgacggcga gatcattggg ctgtcgggtct tcaaacagga ggacggcccc	7260
	aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga	7320
15	ggggtcgccg gtatgtctgt gcgggcggtt ccggcggggtt tattgtctgt gatgatcgtc	7380
	cgacagattc caacgggaat ctgggtggatg cgcattctca tcctcggcgc acttaattatt	7440
20	tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgcggggcgg ggtcgcggcg	7500
	acggtaggcg ctgtgcagcc gctgatggc gtgttcattct ctgccgctct gctaggtagc	7560
	ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgccccgt ggcgctgttg	7620
25	gtgttgacac caaacgcagc gctagatcct gtccggcgctc cagcgggcct ggcggggcg	7680
	gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc	7740
30	acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agcttttagtg	7800
	tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc	7860
	ctgatcggag cgggtttaac ctacttcctt tgggtccggg ggatctcgcg actcgaacct	7920
35	acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg	7980
	catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcgggtg agcaatggat	8040
40	aggggagttg atatcgtaaa cgttcacttc taaagaaata gcgccactca gcttcctcag	8100
	cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca	8160
	cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata	8220
45	tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga	8280
	tcattccgtgt ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat	8340
50	gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct	8400
	gcctgtatcg agtggtgatt ttgtgccgag ctgccggctg gggagctgtt ggctggctgg	8460
	tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg	8520
55	gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat	8580
	tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca	8640
60	gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa	8700
	agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa	8760

gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820
 tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880
 5 ccctaaaggg agccccgat tttagagcttg acgggggaaag ccggcgaaacg tggcgagaaa 8940
 ggaagggaag aaagcgaaag gagcggggcgc cattcaggct gcgcaactgt tgggaagggc 9000
 10 gatcgggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggggatgt gctgcaaggc 9060
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120
 aattaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca 9180
 15 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300
 20 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420
 tcggggcgcg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480
 25 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600
 30 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaaac 9660
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgac gacaagaccg 9720
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaattgggcag 9780
 35 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
 40 tcccttcccg cttcagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960
 agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggctcggtc 10020
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080
 45 ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140
 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200
 50 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260
 agtggagcat ttttgacaag aaatatattgc tagctgatag tgaccttagg cgacttttga 10320
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcgggc 10380
 55 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc 10440
 gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgac ttgatccct 10500
 60 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560
 cttaccagag ggcgccccag ctggcaattc cgttcgctt gctgtccata aaaccgcccc 10620

gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggtt 10680
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
 5 actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcgga 10800
 gcgtgaagct tgcatgcctg caggctcgacg gcgcgccgag ctctctgagc aaatttacac 10860
 10 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920
 tttgatattgc gataaatttt tatatttggt actaaattta taacaccttt tatgctaacg 10980
 tttgccaaca cttagcaatt tgcaagttga ttaattgatt cttaaattatt tttgtcttct 11040
 15 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
 20 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280
 taccacacac aagttttgag gtgcatgcat ggatgccttg tggaaagttt aaaaatatatt 11340
 25 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttg aggatgcaat 11400
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460
 30 ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat 11520
 agccagccca ccgcggtggg cgccgcctg cagtctagaa ggcctcctgc tttaatgaga 11580
 tatgcgagac gcctatgac gcctgatatt tgctttcaat tctgttgtgc acgttgtaaa 11640
 35 aaacctgagc atgtgtagct cagatcctta ccgcgggtt cggttcattc taatgaatat 11700
 atcaccggt actatcgat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760
 40 gtcgacgaat tcgagctcg gcgcctcta gaggatcgat gaattcagat cggctgagtg 11820
 gctccttcaa cggtgcggtt ctgtcagttc caaacgtaaa acggcttgtc ccgcgtcatc 11880
 ggccgggggtc ataacgtgac tcccttaatt ctccgctcat gatcagattg tcgtttcccg 11940
 45 ccttcagttt aaactatcag tgtttgacag gatatttg cggttaaacc taagagaaaa 12000
 gagcgtttat tagaataatc ggatatttaa aaggcgtga aaaggtttat ccttcgtcca 12060
 50 tttgtatgtg catgccaacc acagggttcc cca 12093

<210> 28

<211> 12085

<212> DNA

55 <213> artificial sequence

<220>

<221> misc_feature

60 <223> pflanzlicher Expressionsvektor mit einer Promotor-Terminator-Expr
 essionsskassette

	<400> 28	
	gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaaacg atccgacagc	60
5	gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcca	120
	tagtgggagg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc	180
	ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt	240
10	atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga	300
	ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaaa gtgtcaagca	360
15	tgacaaagtt gcagccgaat acagtgatcc gtgcgcgcct ggacctgttg aacgaggtcg	420
	gcgtagacgg tctgacgaca cgcaaaactgg cggaacgggt ggggggttcag cagccggcgc	480
	tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg	540
20	cggagaatca tacgcattcg gtgcgagag ccgacgacga ctggcgctca tttctgatcg	600
	ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg	660
25	ccggcacgcg accggggcga ccgcagatgg aaacggccga cgcgcagctt cgcttcctct	720
	gcgaggcggg tttttcggcc ggggacggcg tcaatgcgct gatgacaatc agctacttca	780
	ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga	840
30	ccgttgaaca ggctccgctc tcgccgctgt tgcgggcccgc gatagacgcc ttcgacgaag	900
	ccggtccgga cgcagcgctc gagcaggac tcgcggtgat tgcgatgga ttggcgaaaa	960
35	ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc	1020
	tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt	1080
	ccaccgcgct agacgcccg agcagcccgc tacgggcttt tcatgccct gccctagcgt	1140
40	ccaagcctca cggccgcgct cggcctctct ggccggccttc tggcgctctt ccgcttcctc	1200
	gctcactgac tcgctgcgct cggctcgttcg gctgcggcga gcggtatcag ctactcaaa	1260
45	ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	1320
	aggccagcaa aaggccagga accgtaaaaa ggccgpgttg ctggcgcttt tccataggct	1380
	ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgcgac	1440
50	aggactataa agataccagg cgtttccccc tggaagctcc ctcgctcgct ctctgttcc	1500
	gacctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt	1560
55	ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt	1620
	tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac	1680
	ggcgtcagcc gggcaggata ggtgaagtag gccaccgcg gagcgggtgt tccttcttca	1740
60	ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctcgga ggctggccgg	1800
	ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga	1860

	agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa	1920
5	aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca	1980
	aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc	2040
	tgggcccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt	2100
10	tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg	2160
	gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa	2220
15	aacggccggg ggggtgcgct gattgccaa gacgtcccca tgcgctccat caagaagagc	2280
	gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc	2340
	gacgctcacc gggctgggtg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa	2400
20	cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata	2460
	cctcgcgaa aacttgccc tctactgacag atgaggggcg gacgttgaca cttgaggggc	2520
25	cgactcacc ggcgcgcgct tgacagatga ggggcaggct cgatttcggc cgcgacgtg	2580
	gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgcgagtt tccacagat	2640
	gatgtggaca agcctgggga taagtgcct gcggtattga cacttgagg gcgcgactac	2700
30	tgacagatga gggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc	2760
	gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt	2820
35	ccgcccgttt ttccggccacc gctaacctgt cttttaacct gcttttaaac caatatttat	2880
	aaaccttggt tttaaccagg gctgcgcct gtgcgcgtga ccgcgcacgc cgaagggggg	2940
	tgccccccct tctcgaacct tccgggccc ctaacgcggg cctcccatcc ccccaggggc	3000
40	tgcgcccctc ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca	3060
	ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca	3120
45	ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg	3180
	gcggcctggg tggcggcctg cccttcactt cggcggtcgg ggcattcacg gacttcatgg	3240
	cggggcccgc aatttttacc ttgggcattc ttggcatagt ggtcgcggtt gccgtgctcg	3300
50	tgttcggggg tgcgataaac ccagcgaacc atttgagggtg ataggtaaga ttataccgag	3360
	gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag	3420
55	ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa	3480
	tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg	3540
	ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaaact	3600
60	tgcatggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata	3660
	attgggtaat gactocaact tattgatagt gttttatggt cagataatgc ccgatgactt	3720

	tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgccccagc cgtgccagggt	3780
5	gctgcctcag attcagggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt	3840
	gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca	3900
	cgtcaaaggg tgacagcagg ctcataagac gccccagcgt cgccatagtg cgttcaccga	3960
10	atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc	4020
	gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgagacg atgacgtcac	4080
15	tgcccggtg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat	4140
	cgtgttgagg ccaacgcca taatgcgggc tgttgccgg catccaacgc cattcatggc	4200
	catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg	4260
20	ccatgtttta cggcagttag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt	4320
	acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg	4380
25	agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg	4440
	tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga	4500
	aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt	4560
30	cttggtataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa	4620
	taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc	4680
35	gtaaaagata cggaaggaat gtctcctgct aagggtatata agctgggtggg agaaaatgaa	4740
	aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg	4800
	gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaagggt cctgcacttt	4860
40	gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg	4920
	gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac	4980
45	aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc	5040
	ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg	5100
	gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag	5160
50	cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa	5220
	gatggcaaag taagtggctt tattgatctt gggagaagcg gcaggcgga caagtggat	5280
55	gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag	5340
	ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta	5400
	ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg	5460
60	caccgacttc ttccgcatca agtggttttg ctctcaggcc gagggccacg gcaagtat	5520
	gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga	5580

	cggccagacg gtctacggga cggacttcat tgccgataag gtggattatc tggacaccaa	5640
5	ggcaccaggc ggggtcaaatac aggaataagg gcacattgcc ccggcgtgag tcggggcaat	5700
	cccgcgaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat	5760
	cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc	5820
10	gccccgcgaa accttccagt ccgtcggctc gatgggtccag caagctacgg ccaagatcga	5880
	gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatacggccg ccgtggagcg	5940
15	ttcgcgtcgt ctgcaacagg aggcggcagg tttggcgaa tcgatgacca tcgacacgcg	6000
	aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggctcag	6060
	cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct	6120
20	ttccttggtc gatattgcgc cgtggccgga cagcatgcga gcgatgccaa acgacacggc	6180
	ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa	6240
25	ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgctc agctgcgggc	6300
	cgacgatgac gaactgggtg ggcagcaggt gttggagtac gcgaagcgca cccctatcgg	6360
	cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg	6420
30	ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt	6480
	cacgtccgac cgcgttgggc acctggaatc ggtgtcgtc ctgcaccgct tccgcgtcct	6540
35	ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct	6600
	gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac	6660
	ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga	6720
40	aaccttcgc ctcatgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt	6780
	cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct ggggtcaatga	6840
45	tgacctggtg cattgcaaac gctagggcct tgtggggta gttccggctg ggggttcagc	6900
	agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc	6960
	tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa	7020
50	ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc	7080
	cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tggtcgggtc cgtttacgag	7140
55	cacgaggaga aaaagcccat ggaggcgctt gctgaacggg tgcgagatgc cgtggcattc	7200
	ggcgcctaca tcgacggcga gatcattggg ctgtcgggtc tcaaacagga ggacggcccc	7260
	aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga	7320
60	ggggtcggcg gtatgctgct gcgggcgttg cggcggggtt tattgctcgt gatgatcgtc	7380
	cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaattatt	7440

	tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccggggcg ggtcgggcg	7500
5	acggtagggc ctgtgcagcc gctgatgggc gtgttcatct ctgccgctct gctaggtagc	7560
	ccgatacgat tgatggcggt cctgggggct atttgcgaa ctgccccgt ggcgctgttg	7620
	gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggcggggcg	7680
10	gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc	7740
	acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg	7800
15	tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc	7860
	ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct	7920
	acagttgttt ccttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg	7980
20	catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat	8040
	aggggagttg atatcgtaaa cgttcacttc taaagaaata gcgccactca gcttctcag	8100
25	cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca	8160
	cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata	8220
	tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga	8280
30	tcacccgtgt ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat	8340
	gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct	8400
35	gcctgtatcg agtggtgatt ttgtgccgag ctgccggctc gggagctgtt ggctggctgg	8460
	tggcaggata tatttggttg taaacaaatt gacgcttaga caacttaata acacattgcg	8520
	gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat	8580
40	tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca	8640
	gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa	8700
45	agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa	8760
	gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg	8820
	tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa	8880
50	ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaaac tggcgagaaa	8940
	ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc	9000
55	gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc	9060
	gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg	9120
	aattaaattcc catcttgaaa gaaatatagt ttaaataatt attgataaaa taacaagtca	9180
60	ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt	9240
	tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta	9300

5 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420
 tcggggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480
 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540
 10 ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca 9600
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac 9660
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgac gacaagaccg 9720
 15 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840
 20 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900
 tcccttcccg cttcagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960
 agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggtcggtc 10020
 25 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080
 ccgattgtct gttgtgcccc gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140
 30 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320
 35 acgcgcaata atgggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcgggc 10380
 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc 10440
 40 gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gtcctatgac ttgatccct 10500
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560
 45 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt 10680
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
 50 actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcgga 10800
 gcgtgaagct tgcctgcctg caggtcgacg gcgcgcgag ctcctcgagc aaatttacac 10860
 55 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980
 tttgccaaca cttagcaatt tgcaagtga ttaattgatt ctaaattatt tttgtcttct 11040
 60 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160

	ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga	11220
5	tttgagggtgc atgaacgtca cgtgggacaaa aggttttagta atttttcaag acaacaatgt	11280
	taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt	11340
	tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat	11400
10	aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga	11460
	ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat	11520
15	agccagcgga tccgatatcg ggcccgcctag cgtaaacctt gctttaatga gatatgcgag	11580
	acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga	11640
	gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaataaat atatcaccgg	11700
20	ttactatcgt atttttatga ataataattct ccgttcaatt tactgattgt ccgtcgacga	11760
	attcgagctc ggccgcgcctc tagaggatcg atgaattcag atcggctgag tggctccttc	11820
25	aacgttgccg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg	11880
	tcataacgtg actcccttaa ttctccgctc atgatcagat tgtcgtttcc cgccttcagt	11940
	ttaaactatc agtgtttgac aggatatatt ggccgggtaaa cctaagagaa aagagcggtt	12000
30	attagaataa tcggatatatt aaaagggcgt gaaaaggttt atccttcgtc catttgatg	12060
	tgcattgcaa ccacaggggt cccca	12085
35	<210> 29	
	<211> 12079	
	<212> DNA	
	<213> artificial sequence	
40	<400> 29	
	gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc	60
	gcgcccagca caggtgacga ggcaaatgac accaacgcat acagcgccag cagaatgcc	120
45	tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc	180
	ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt	240
50	atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga	300
	ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca	360
	tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg	420
55	gcgtagacgg totgacgaca cgcaaaactg cggaacggtt ggggggttcag cagccggcgc	480
	tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg	540
60	cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg	600
	ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg	660

	ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgagctt cgcttcctct	720
	gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca	780
5	ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca	840
	ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag	900
10	ccggtccgga cgcagcgctc gagcaggac tcgcggtgat tgcgatgga ttggcgaaaa	960
	ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc	1020
	tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt	1080
15	ccaccgctc agacgcccg agcagccgc tacgggcttt ttcatgccct gccctagcgt	1140
	ccaagcctca cgcccgcgct cggcctctct ggcgcccttc tggcgctctt ccgcttcctc	1200
20	gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctactcaaa	1260
	ggcggtataa cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	1320
	aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	1380
25	ccgccccct gacgagcatc aaaaaaatcg acgctcaagt cagaggtggc gaaaccgcac	1440
	aggactataa agataccagg cgtttcccc tgggaagctcc ctctgcgct ctctgtttcc	1500
30	gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt	1560
	ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt	1620
	tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac	1680
35	ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tcctttcttca	1740
	ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg	1800
40	ctaccgccgg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga	1860
	agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa	1920
	agggcgccgc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggtaca	1980
45	aatcacggg cgtcgtggac tatgagcag tccgcgagct ggcccgcatc aatggcgacc	2040
	tgggcccctt gggcgccctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt	2100
50	tcggtgatgc cagatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg	2160
	gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa	2220
	aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc	2280
55	gacttcgagg agctgggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc	2340
	gacgctcacc gggctggttg ccctgcgcgc tgggctggcg gccgtctatg gccctgcaaa	2400
60	cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcgggc gttgtggata	2460
	cctcgcggaa aacttggccc tcaactgacag atgaggggag gacgttgaca cttgaggggc	2520

	cgactcacccc	ggcgcggcgt	tgacagatga	ggggcaggct	cgatttcggc	cggcgacgtg	2580
	gagctggcca	gcctcgcaaa	tcggcgaaaa	cgcctgattt	tacgcgagtt	tcccacagat	2640
5	gatgtggaca	agcctgggga	taagtgcctt	gcggtattga	cacttgaggg	gcgcgactac	2700
	tgacagatga	ggggcgcgat	ccttgacact	tgaggggcag	agtgcctgaca	gatgagggggc	2760
10	gcacctattg	acatttgagg	ggctgtccac	aggcagaaaa	tccagcattt	gcaaggggtt	2820
	ccgcccgttt	ttcggccacc	gctaactgtt	cttttaacct	gcttttaaac	caatatttat	2880
	aaaccttggt	tttaaccagg	gctgcgccct	gtgcgcgtga	ccgcgcacgc	cgaagggggg	2940
15	tgccccccct	tctcgaaccc	tcccggcccc	ctaacgcggg	cctcccatcc	ccccaggggc	3000
	tgcccccctc	ggccgcgaac	ggcctcaccc	caaaaatggc	agcgcctggca	gtccttgcca	3060
20	ttgccgggat	cggggcagta	acgggatggg	cgatcagccc	gagcgcgacg	cccgaagca	3120
	ttgacgtgcc	gcaggtgctg	gcatcgacat	tcagcgacca	ggtgccgggc	agtgagggcg	3180
	gcggcctggg	tggcggcctg	cccttcactt	cggccgtcgg	ggcattcacg	gacttcattg	3240
25	cggggccggc	aatttttacc	ttgggcattc	ttggcatagt	ggtcgcgggt	gccgtgctcg	3300
	tgttcggggg	tgcgataaac	ccagcgaacc	atttgagggt	ataggtaaga	ttataccgag	3360
30	gtatgaaaac	gagaattgga	cctttacaga	attactctat	gaagcgccat	atttaaaaag	3420
	ctaccaagac	gaagaggatg	aagaggatga	ggaggcagat	tgccctgaat	atattgacaa	3480
	tactgataag	ataatatatc	ttttatatag	aagatatcgc	cgtatgtaag	gatttcaggg	3540
35	ggcaaggcat	aggcagcgcg	cttatcaata	tatctataga	atgggcaaag	cataaaaact	3600
	tgcatggact	aatgcttgaa	accaggaca	ataaccttat	agcttgtaaa	ttctatcata	3660
40	attgggtaat	gactccaact	tattgatagt	gttttatgtt	cagataatgc	ccgatgactt	3720
	tgtcatgcag	ctccaccgat	tttgagaacg	acagcgactt	ccgtcccagc	cgtgccagggt	3780
	gctgcctcag	attcagggtta	tgccgctcaa	ttcgctgcgt	atatcgcttg	ctgattacgt	3840
45	gcagctttcc	cttcaggcgg	gattcatata	gcggccagcc	atccgtcacc	catatcacca	3900
	cgtcaaaggg	tgacagcagg	ctcataagac	gccccagcgt	cgccatagtg	cgttcaccga	3960
50	atacgtgcgc	aacaaccgtc	ttccggagac	tgtcatacgc	gtaaaacagc	cagcgctggc	4020
	gcgatttagc	cccacatag	cccactgtt	cgtccatttc	cgcgcagacg	atgacgtcac	4080
	tgcccggtg	tatgcgcgag	gttaccgact	gcggcctgag	ttttttaagt	gacgtaaaat	4140
55	cgtgttgagg	ccaacgcccc	taatgcgggc	tgttgcccgg	catccaacgc	cattcatggc	4200
	catatcaatg	atcttctggt	gcgtaccggg	ttgagaagcg	gtgtaagtga	actgcagttg	4260
60	ccatgtttta	cggcagtgag	agcagagata	gcgctgatgt	ccggcggtgc	ttttgccgtt	4320
	acgcaccacc	ccgtcagtag	ctgaacagga	gggacagctg	atagacacag	aagccactgg	4380

	agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg	4440
	tggtttcaaa atcgggtccg tcgatactat gttatacgcc aactttgaaa acaactttga	4500
5	aaaagctggt ttctggtatt taagggtttta gaatgcaagg aacagtgaat tggagttcgt	4560
	cttgtttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa	4620
10	taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc	4680
	gtaaaagata cggaaggaat gtctcctgct aagggtatata agctgggtggg agaaaatgaa	4740
	aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg	4800
15	gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt	4860
	gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg	4920
20	gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac	4980
	aggctcttct actccatcga catatcggat tgtccctata cgaatagctt agacagccgc	5040
	ttagcdgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg	5100
25	gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag	5160
	cccgaagagg aacttgtctt ttccacggc gacctgggag acagcaacat ctttgtgaaa	5220
30	gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat	5280
	gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag	5340
	ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatTTTa	5400
35	ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg	5460
	caccgacttc ttccgcatca agtgTTTTg ctctcaggcc gaggccacg gcaagtattt	5520
40	gggcaagggg tcgctggat tcgtgcagg caagattcgg aataccaagt acgagaagga	5580
	cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa	5640
	ggcaccaggc ggggtcaaac aggaataagg gcacattgcc ccggcgtgag tccgggcaat	5700
45	cccgcaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat	5760
	cgacgggggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc	5820
50	gccccgcgaa accttccagt ccgtcggctc gatgggtccag caagctacgg ccaagatcga	5880
	gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccacggcccg ccgtggagcg	5940
	ttcgcgtcgt ctgcaacagg aggcggcagg tttggcgaaag tcgatgacca tcgacacgcg	6000
55	aggaactatg acgaecaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag	6060
	cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct	6120
60	ttccttgctc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc	6180
	ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa	6240

	ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgctcg agctgcgggc	6300
	cgacgatgac gaactggtgt ggcagcaggt gttggagtag gcgaagcgca cccctatcgg	6360
5	cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg	6420
	ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt	6480
10	cacgtccgac cgcgttgggc acctggaatc ggtgtcgtcg ctgcaccgct tccgcgtcct	6540
	ggaccgtggc aagaaaacgt cccgttgcca ggtcctgata gacgaggaaa tcgtcgtgct	6600
	gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgcggac	6660
15	ggcccagcgg atgttcgact atttcagctc gcaccgggag ccgtaccgcg tcaagctgga	6720
	aaccttccgc ctcatgtgcg gatcggatc caccgcgctg aagaagtggc gcgagcaggt	6780
20	gggcgaagcc tgcgaagagt tgcgaggcag cggcctgggtg gaacacgcct ggggtcaatga	6840
	tgacctggtg cattgcaaac gctagggcct tgtggggctca gttccggctg ggggttcagc	6900
	agccagpgct ttactggcat ttcaggaaca agcgggcact gtcgacgca cttgcttcgc	6960
25	tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa	7020
	ttgacattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc	7080
30	cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tggtcgggtc cgtttacgag	7140
	cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc	7200
	ggcgcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc	7260
35	aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaaca gcgaggccga	7320
	ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc	7380
40	cgacagattc caacgggaat ctggtggatg cgcattctca tctcggcgc acttaatt	7440
	tcgtattctt ggagcttggt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg	7500
	acggtaggcg ctgtgcagcc gctgatggtc gtgttcattc ctgccgctct gctaggtagc	7560
45	ccgatacgat tgatggcggc cctgggggct atttgcggaa ctgcgggcgt ggcgctggtg	7620
	gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggcggggcg	7680
50	gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc	7740
	acctttaccg cctggcaact ggcgcccgga ggacttctgc tcgttccagt agcttttagtg	7800
	tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc	7860
55	ctgatcggag cgggtttaac ctacttcctt tgggtccggg ggatctcgcg actcgaacct	7920
	acagttgttt cttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg	7980
60	catcaggccg acagtccgaa cttegggtcc ccgacctgta ccattcggtg agcaatggat	8040
	aggggagttg atatcgtaa cgttcacttc taaagaaata gcgccactca gcttcctcag	8100

	cggttttatac cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca	8160
	cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata	8220
5	tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga	8280
	tcattcctgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat	8340
10	gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct	8400
	gcctgtatcg agtgggtgatt ttgtgcgcgag ctgccggctcg gggagctggt ggctggctgg	8460
	tggcaggata tattgtgggtg taaacaaatt gacgcttaga caacttaata acacattgcy	8520
15	gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat	8580
	tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca	8640
20	gcaggcgaaa atcctgtttg atggtgggtc cgaaatcggc aaaatccctt ataaatcaaa	8700
	agaatagccc gagatagggg tgagtgttgt tccagtttgg aacaagagtc cactattaaa	8760
	gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg	8820
25	tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa	8880
	ccctaagggg agccccgat ttagagcttg acggggaaaag ccggcgaaacg tggcgagaaa	8940
30	ggaagggaag aaagcgaaaag gagcgggagc cattcaggct gcgcaactgt tgggaagggc	9000
	gatcgggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggggatgt gctgcaaggc	9060
	gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg	9120
35	aattaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca	9180
	ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt	9240
40	tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta	9300
	tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggagc cgtcgaagct	9360
	agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa	9420
45	tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc gccaaagtct	9480
	tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg	9540
50	ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca	9600
	tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaa	9660
	agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg	9720
55	gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaatgggcag	9780
	gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg	9840
60	gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgccccaa tagcagccag	9900
	tcccttcccg cttcagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc	9960

agccacgata gccgcgctgc ctgcctctgc agttcattca gggcaccgga caggtcgggtc 10020
ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080
5 ccgattgtct gttgtgcccc gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140
cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200
10 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260
agtggagcat ttttgacaag aatatattgc tagctgatag tgaccttagg cgacttttga 10320
acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcgggc 10380
15 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgcg 10440
gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gtcctgatc ttgatccct 10500
gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560
20 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620
gtctagptat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt 10680
25 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
actgggtttc tacgtgttcc gtttcttta gcagccttg cgccctgagt gcttgcgga 10800
gcgtgaagct tgcatgcctg caggctgacg gcgcgccgag ctctcgagc aaatttacac 10860
30 attgccacta aacgtctaaa cccttgtaatt ttgtttttgt ttactatgt gtgttatgta 10920
tttgatttgc gataaatttt tatatttgggt actaaattta taacaccttt tatgctaacg 10980
35 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040
aaatacatat actaatcaac tggaaatgta aatatattgt aatatttcta ctataggaga 11100
attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160
40 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280
45 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340
tggaaatgat ttgcatgga gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400
aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataagagga 11460
50 ttttgcaata ctttcattca tacacactca ctaagtttta cagattata atttcttcat 11520
agccagcaga tctgcccggc tcgatcccg gccatggcct gctttaatga gatatgcgag 11580
55 acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga 11640
gcgtgtgtag ctcagatcct taccgccggt ttcggttcat tctaataat atatcaccg 11700
ttactatcgt atttttatga ataatttct ccgttcaatt tactgattgt ccgtcgacga 11760
60 gctcggcgcg cctctagagg atcgatgaat tcagatcggc tgagtggctc cttcaacggt 11820

gcggttctgt cagttccaaa cgtaaaacgg cttgtcccgc gtcacggcg ggggtcataa 11880
 cgtgaactccc ttaattctcc gctcatgac agattgtcgt ttcccgccct cagtttaaac 11940
 5 tatcagtgtt tgacaggata tattggcggg taaacctaa agaaaagagc gtttattaga 12000
 ataattggat atttaaaagg gcgtgaaaag gtttatccct cgtccatttg tatgtgcatg 12060
 ccaaccacag ggttcccca 12079
 10
 <210> 30
 <211> 13002
 <212> DNA
 15 <213> artificial sequence
 <220>
 <223> pflanzlicher Expressionsvektor mit zwei Promotor-Terminator-Expre
 ssionskassetten
 20
 <400> 30
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60
 gcgcccagca cagggtgcga ggcaaattgc accaacgcat acagcgccag cagaatgcca 120
 25 tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180
 ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240
 30 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360
 tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420
 35 gcgtagacgg tctgacgaca cgaaactgg cggaacgggt ggggggttcag cagccggcgc 480
 tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540
 40 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660
 45 ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720
 gcgaggcggg ttttctggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca 840
 50 ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc ttcgacgaag 900
 ccggtccgga cgcagcgttc gagcaggac tcgcggtgat tgcgatgga ttggcgaaaa 960
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020
 55 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080
 ccacdcgctc agacgcccg agcagccgc tacgggcttt ttcatgccct gccctagcgt 1140
 60 ccaagcctca cggccgcgct cggcctctct ggccgcttc tggcgctctt ccgcttctc 1200
 gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcaactdaaa 1260

	ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	1320
5	aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	1380
	ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac	1440
	aggactataa agataccagg cgtttccccc tggaagctcc ctctgctgct ctctgtttcc	1500
10	gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt	1560
	ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt	1620
15	tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttcttggt tgtatccaac	1680
	ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca	1740
	ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg	1800
20	ctaccgccgg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga	1860
	agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa	1920
25	agggcgccgc gggcgccatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca	1980
	aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc	2040
	tgggcccgcct gggcgccctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt	2100
30	tcggtgatgc cagcatctc gccctgctgg cgaagatcga agagaagcag gacgagcttg	2160
	gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa	2220
35	aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc	2280
	gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga ggcgctttgc	2340
	gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa	2400
40	cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata	2460
	cctcgcgga aacttgccc tctactgacag atgaggggag gacgttgaca cttgaggggc	2520
45	cgactacccc ggcgcgccgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg	2580
	gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat	2640
	gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac	2700
50	tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc	2760
	gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt	2820
55	ccgcccgttt ttccggccacc gctaacctgt cttttaacct gcttttaaac caatatattat	2880
	aaaccttggt ttttaaccagg gctgcgcctt gtgcgcgtga ccgcgcacgc cgaagggggg	2940
	tgccc'ccct tctcgaaccc tcccgcccg ctaacgcggg cctcccatcc ccccaggggc	3000
60	tgcgcccctc ggccgcgaac ggcctcacc caaaaatggc agcgtggca gtccttgcca	3060
	ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca	3120

	ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg	3180
5	gccccctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcatgg	3240
	cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg	3300
	tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag	3360
10	gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag	3420
	ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa	3480
15	tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg	3540
	ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact	3600
	tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata	3660
20	attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt	3720
	tgtcatgcag ctccaccgat ttgagaacg acagcgactt ccgtcccagc cgtgccagggt	3780
25	gctgcctcag attcagggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt	3840
	gcagctttcc cttcaggcgg gattcatata gggccagcc atccgtcatc catatcacca	3900
	cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga	3960
30	atacgtgcgc aacaaccgtc ttccggagac tgtcatagc gtaaacagc cagcgctggc	4020
	gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac	4080
35	tgcccggtg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat	4140
	cgtgttgagg ccaacgccc taatgcgggc tgttgcccg catccaacgc cattcatggc	4200
	catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg	4260
40	ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt	4320
	acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg	4380
45	agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg	4440
	tggtttcaaa atcggctccg tcgatactat gttatagcc aactttgaaa acaactttga	4500
	aaaagctggt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt	4560
50	cttggtataa ttagcttctt ggggtatctt taaatctgt agaaaagagg aaggaaataa	4620
	taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc	4680
55	gtaaaagata cggaaggat gtctcctgct aaggatatata agctggtggg agaaaatgaa	4740
	aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg	4800
	gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaagggt cctgcacttt	4860
60	gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg	4920
	gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac	4980

aggctctttc actccatoga catatcggat tgtccctata cgaatagctt agacagccgc 5040
5 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100
gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag 5160
cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220
10 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280
gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340
ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400
15 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460
caccgacttc ttccgcatca agtgttttgg ctctcaggcc gagggccacg gcaagtattt 5520
20 gggcaagggg tcgctggat tctgtcaggg caagattcgg aataccaagt acgagaagga 5580
cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
ggcaccaggc ggggtcaaata aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700
25 cccgcaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760
cgacgcgggg ttttccgccc aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820
30 gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatoga 5880
gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccacggccg ccgtggagcg 5940
ttcgcgtcgt ctgcaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000
35 aggaactatg acgaccaaga agcgaataac cgccggcgag gacctggcaa aacaggtcag 6060
cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aatgcagct 6120
40 ttccttggtc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc 6180
ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300
45 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360
cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420
50 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480
cacgtccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgcgtcct 6540
ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600
55 gtttgcgtggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgctgccgac 6660
ggccdgacgg atgttcgact atttcagtc gcaccgggag ccgtaccgc tcaagctgga 6720
60 aaccttccgc ctcatgtgcg gatcggattc cccccgctg aagaagtggc gcgagcaggt 6780
cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct ggggtcaatga 6840

	tgacctggtg cattgcaaac gctagggcct tgtgggggtca gttccggctg ggggttcagc	6900
5	agccagcgct ttactggcat ttcaggaaca agcgggact gctcgacgca cttgcttcgc	6960
	tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa	7020
	ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc	7080
10	cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag	7140
	cacgaggaga aaaagcccat ggaggcgctc gctgaacggt tgcgagatgc cgtggcattc	7200
15	ggcgcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc	7260
	aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga	7320
	ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc	7380
20	cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatat	7440
	tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg	7500
25	acggtaggcg ctgtgcagcc gctgatggc gtgttcattc ctgccgtctc gctaggtagc	7560
	ccgatacgat tgatggcggt cctgggggct atttgcgga ctgcgggcgt ggcgctgttg	7620
	gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggcggggcg	7680
30	gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctccgt gcctctgctc	7740
	acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttcagat agctttagt	7800
35	tttgatccgc caatccgat gcctacagga accaatgttc tcggcctggc gtggctcggc	7860
	ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct	7920
	acagttgttt cttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg	7980
40	catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcgggt agcaatggat	8040
	aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag	8100
45	cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca	8160
	cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctcgag ggagatgata	8220
	tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga	8280
50	tcattccgtg ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat	8340
	gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct	8400
55	gcctgtatcg agtgggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg	8460
	tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattg	8520
	gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat	8580
60	tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca	8640
	gcaggcgaaa atcctgtttg atgggtggtc cgaaatcggc aaaatccctt ataaatcaaa	8700

	agaatagccc gagatagggg tgagtgttgt tccagtttgg aacaagagtc cactattaaa	8760
5	gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg	8820
	tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa	8880
	ccctaaaggg agccccgat ttagagcttg acgggggaaag cgggcgaacg tggcgagaaa	8940
10	ggaaggggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc	9000
	gatcgggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc	9060
15	gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagt	9120
	aatttaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca	9180
	ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt	9240
20	tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta	9300
	tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct	9360
25	agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa	9420
	tcgggagcgg cgataccgta aagcacgagg aagcggctcag cccattcgcc gccaaagctct	9480
	tcagcaatat cacgggtagc caacgctatg tctgatagc ggtccgccac acccagccgg	9540
30	ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca	9600
	tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaa	9660
35	agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg	9720
	gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaatgggcag	9780
	gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg	9840
40	gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag	9900
	tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc	9960
45	agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggtcggtc	10020
	ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag	10080
	ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc ggccggagaa	10140
50	cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga	10200
	tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc	10260
55	agtggagcat ttttgacaag aaatatattgc tagctgatag tgaccttagg cgacttttga	10320
	acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc	10380
	tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc	10440
60	gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatccct	10500
	gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac	10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620
gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt 10680
5 ttccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
actggctttc tacgtgttcc gcttccttta gcagcccttg cgcctgagt gcttgcgga 10800
10 gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctctcgagc aaatttacac 10860
attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920
15 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980
tttgccaaca cttagcaatt tgcaagttga ttaattgatt cttaaattatt tttgtcttct 11040
aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100
20 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160
ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
25 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280
taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340
tggaaatgat ttgcatggaa gccatgtgta aaacctgac atccacttgg aggatgcaat 11400
30 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460
ttttgcaata ctttcattca tacacactca ctaagtttta cagattata atttcttcat 11520
35 agccagccca ccgcggtggg cggccgcctg cagtctagaa ggctcctgc tttaatgaga 11580
tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 11640
aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat 11700
40 atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760
gtcgagcaaa tttacacatt gccactaaac gtctaaaccc ttgtaatttg tttttgtttt 11820
45 actatgtgtg ttatgtattt gatttgcgat aaatttttat atttggtact aaatttataa 11880
caccttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 11940
aattattttt gtcttctaaa tacatatact aatcaactgg aatgtaaat atttgcta 12000
50 atttctacta taggagaatt aaagtgagt aatatggtac cacaaggttt ggagatttaa 12060
ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 12120
aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt 12180
55 tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgccctgtgg 12240
aaagttaaaa aatatttttg aatgatttg catggaagcc atgtgtaaaa ccatgacatc 12300
60 cacttggagg atgcaataat gaagaaaact acaaatttac atgcaactag ttatgcatgt 12360
agtctatata atgaggattt tgcaatactt tcattcatac acactcacta agttttacac 12420

gattataatt tcttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct 12480
 5 ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttggtgca 12540
 cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccgggttc gggttcattct 12600
 aatgaatata tcaccggtta ctatcgattt tttatgaata atattctccg ttcaatttac 12660
 10 tgattgtccg tcgacgaatt cgagctcggc gcgcctctag aggatcgatg aattcagatc 12720
 ggctgagtgg ctccctcaac gttgcgggtc tgtcagttcc aaacgtaaaa cggcttggtcc 12780
 cgcgctcatcg gcgggggtca taacgtgact cccttaattc tccgctcatg atcagattgt 12840
 15 cgtttcccg cttcagttta aactatcagt gtttgacagg atatattggc gggtaaacct 12900
 aagagaaaag agcgtttatt agaataatcg gatatttaaa agggcgtgaa aagggtttatc 12960
 20 ctctgctccat ttgtatgtgc atgccaacca cagggttccc ca 13002

 <210> 31
 <211> 13905
 25 <212> DNA
 <213> artificial sequence

 <220>
 30 <223> pflanzlicher Expressionsvektor mit drei Promotor-Terminator-Expre
 ssionskassetten

 <400> 31
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60
 35 gcgcccagca cagggtgcga ggcaaattgc accaacgcat acagcgccag cagaatgcca 120
 tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180
 40 ataatcaggc cgatgccgac agcgctcgagc gcgacagtgc tcagaattac gatcaggggt 240
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360
 45 tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420
 gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc 480
 50 ttactgggca cttcaggaac aagcgggagc tgctcgacgc actggccgaa gccatgctgg 540
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660
 55 ccggcacgcg accggggcga ccgcagatgg aaacggccga cgcgacgctt cgcttctct 720
 gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780
 60 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca 840
 ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc ttcgacgaag 900

	ccggtccgga cgcagcggttc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa	960
	ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc	1020
5	tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt	1080
	ccaccgcgtc agacgcccgt agcagcccg c tacgggcttt ttcattgcct gccctagcgt	1140
10	ccaagcctca cggccgcgct cggcctctct ggccggccttc tggcgctctt ccgcttcctc	1200
	gctcactgac tcgctgcgct cggtcggttc gctgcggcga gcggtatcag ctactcaaa	1260
	ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	1320
15	aggccagcaa aaggccagga accgtaaaaa ggccgcggtt ctggcgtttt tccataggct	1380
	ccgccccctt gacgagcatc acaaaaaatcg acgctcaagt cagaggtggc gaaacccgac	1440
20	aggactataa agataccagg cgtttcccc tggaaagctcc ctctgctgct ctctgttcc	1500
	gaccctgccc cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt	1560
	ccgctgcata acctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt	1620
25	tcgcacgata tacaggattt tgccaaaggg ttctgtgtaga ctttcttgg tgtatccaac	1680
	ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca	1740
30	ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg	1800
	ctaccgccgg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga	1860
	agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa	1920
35	aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca	1980
	aaatcacggg cgctgtggac tatgagcag tccgcgagct ggcccgcatc aatggcgacc	2040
40	tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt	2100
	tcggtgatgc cagcatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg	2160
	gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa	2220
45	aacggccggg ggggtgcgct gattgccaag cagtcacca tgcgctccat caagaagagc	2280
	gacttcgagg agctgggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc	2340
50	gacgctcacc gggctggttg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa	2400
	cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata	2460
	cctcgccgaa aacttgccc tcactgacag atgagggcg gacgttgaca cttgaggggc	2520
55	cgactacccc ggcgcggcgt tgacagatga ggggcaggct cgatttcggc cggcgacgtg	2580
	gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat	2640
60	gatgtggaca agcctgggga taagtgccct gcggtattga cacttgaggg gcgcgactac	2700
	tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc	2760

gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggttt 2820
 ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880
 5 aaaccttggt ttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg 2940
 tgccccccct tctcgaacct tcccggcccc ctaacgcggg cctcccatcc cccagggggc 3000
 10 tgcgcccctc ggccgcgaac ggctcacc caaaaatggc agcgctggca gtccttgcca 3060
 ttgcggggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgaggggc 3180
 15 gcggcctggg tggcggcctg cccttcactt cggcgcgtcg ggcattcacg gacttcatgg 3240
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300
 20 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360
 gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420
 ctaccagagc gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480
 25 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaaact 3600
 30 tgcattggact aatgcttgaa acccaggaca ataaccctat agcttgtaaa ttctatcata 3660
 attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780
 35 gctgcctcag attcagggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt 3840
 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900
 40 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgtggc 4020
 gcgatttagc ccgacatag cccactgtt cgtccatttc cgcgacagc atgacgtcac 4080
 45 tgcccggtg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140
 cgtgttgagg ccaacgcca taatgcgggc tgttgcccg catccaacgc cattcatggc 4200
 50 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtg ttttgccgtt 4320
 acgcaccacc ccgtcagttag ctgaacagga gggacagctg atagacacag aagccactgg 4380
 55 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg 4440
 tggtttcaaa atcggtccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500
 60 aaaagctgtt ttctgggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620

taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680
 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctgggtggg agaaaatgaa 4740
 5 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860
 10 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980
 aggctctttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040
 15 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag 5160
 20 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340
 25 ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatatta 5400
 ctggatgaat tgtttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460
 30 caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggcccacg gcaagtat 5520
 gggcaagggg tcgctggat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640
 35 ggcaccaggc ggggtcaaac aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700
 cccgcaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760
 40 cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820
 gccccgcgaa accttccagt ccgtcggctc gatgggtccag caagctacgg ccaagatcga 5880
 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940
 45 ttgcgctcgt ctgcaacagg agggcgagc tttggcgaag tcgatgacca tcgacacgcg 6000
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060
 50 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120
 ttcttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgcaa acgacacggc 6180
 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240
 55 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300
 cgacgatgac gaactgggtg ggcagcagg gttggagtac gcgaagcgca cccctatcgg 6360
 60 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggg cgatcaatgg 6420
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480

cacgtccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgcgtcct 6540
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgac gacgaggaaa tcgtcgtgct 6600
 5 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgcgcgac 6660
 ggcccgcacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga 6720
 10 aaccttccgc ctcatgtgcg gatcggatc caccgcgtg aagaagtggc gcgagcaggt 6780
 cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct ggggtcaatga 6840
 tgacctggtg cattgcaaac gctagggcct tgtgggggtca gttccggctg ggggttcagc 6900
 15 agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc 6960
 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020
 20 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140
 cacgaggaga aaaagcccat ggaggcgctc gctgaacggt tgcgagatgc cgtggcattc 7200
 25 ggcgcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc 7260
 aaggacgctc acaaggcgca tctgtccggc gttttcgtg agcccgaaaca gcgaggccga 7320
 30 ggggtcgcg gtatgctgct gggggcgctt cggcggggtt tattgctcgt gatgatcgtc 7380
 cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatat 7440
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggagg ggtcgcggcg 7500
 35 acggtaggcg ctgtgcagcc gctgatggc gtgttcattc ctgccgctc gctaggtagc 7560
 ccgatacgat tgatggcggt cctgggggct atttgcgga ctgcgggctt ggcgctgttg 7620
 40 gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggcgggggag 7680
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccg gcctctgctc 7740
 acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttcagat agcttttagt 7800
 45 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860
 ctgatcggag cgggtttaac ctacttcctt tgggtccggg ggatctcgc actcgaacct 7920
 50 acagttgttt cttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg 7980
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcgggt agcaatggat 8040
 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gtttcctcag 8100
 55 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220
 60 tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280
 tcatccgtgt ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340

	gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccga ctgatgggct	8400
	gcctgtatcg agtgggtgatt ttgtgccgag ctgccggtcg gggagctggt ggctggctgg	8460
5	tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg	8520
	gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat	8580
10	tgcccttcac cgcttgccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca	8640
	gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa	8700
	agaatagccc gagatagggt tgagtgttgt tccagtttg aacaagagtc cactattaaa	8760
15	gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gccactacg	8820
	tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa	8880
20	ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa	8940
	ggaaggggag aaagcgaaa gagcggggcg cattcaggct gcgcaactgt tgggaagggc	9000
	gatcgggtcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc	9060
25	gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg	9120
	aattaattcc catcttgaaa gaaatatagt ttaaataatt attgataaaa taacaagtca	9180
30	ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt	9240
	tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta	9300
	tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct	9360
35	agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa	9420
	tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc gccaaagctct	9480
40	tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac accagccgg	9540
	ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca	9600
	tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac	9660
45	agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg	9720
	gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaatgggcag	9780
50	gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg	9840
	gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag	9900
	tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc	9960
55	agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggtcggtc	10020
	ttgacaaaaa gaaccgggcg ccctgcgct gacagccgga acacggcggc atcagagcag	10080
60	ccgattgtct gttgtgcca gtcatagccg aatagcctct ccacccaagc ggccggagaa	10140
	cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga	10200

tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260
agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320
5 acgcgcaata atgggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380
tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaacgg cttgtcccgc 10440
10 gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgac ttgatccct 10500
gcgccatgag atccttggcg gcaagaagc catccagttt actttgcagg gcttcccaac 10560
cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620
15 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt 10680
ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740
20 actggctttc tacgtgttcc gcttccttta gcagccctg cgccctgagt gcttgcgcca 10800
gcgtgaagct tgcattgctg caggctcgac gcgcgccgag ctccctcgagc aaatttacac 10860
attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920
25 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980
tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040
30 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100
attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160
ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220
35 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280
taccacacac aagttttgag gtgcatgcat ggatgcctg tggaaagttt aaaaatattt 11340
40 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400
aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460
ttttgcaata ctttcattca tacacactca ctaagtttta cagattata atttcttcat 11520
45 agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga 11580
tatgcgagac gcctatgac gcattgatatt tgctttcaat tctgttgtgc acgttgtaaa 11640
50 aaacctgagc atgtgtagct cagatcctta ccgcgggtt cggttcattc taatgaatat 11700
atcacccgtt actatcgat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760
gtcgagcaaa tttacacatt gccactaaac gtctaaacc ttgtaatttg tttttgtttt 11820
55 actatgtgtg ttatgtattt gatttgcgat aaatttttat atttggtact aaatttataa 11880
caqcttttat gctaacgtt gccaacactt agcaatttgc aagttgatta attgattcta 11940
60 aattattttt gtcttctaaa tacatatact aatcaactgg aatgtaaat atttgcta 12000
atttctacta taggagaatt aaagtgagt aatatggtac cacaagggtt ggagatttaa 12060

ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 12120
 aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt 12180
 5 tttcaagaca acaatgttac cacacacaag ttttgagggtg catgcatgga tgccctgtgg 12240
 aaagttaaaa aatatttttg aaatgatttg catggaagcc atgtgtaaaa ccatgacatc 12300
 10 cacttgaggagg atgcaataat gaagaaaact acaaatttac atgcaactag ttatgcatgt 12360
 agtctatata atgaggattt tgcaataact tcattcatatc acactcacta agttttacac 12420
 gattataatt tcttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct 12480
 15 ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttgtgca 12540
 cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccggtttc ggttcattct 12600
 aatgaatata tcacccgta ctatcgtatt tttatgaata atattctccg ttcaatttac 12660
 20 tgattgtccg tcgagcaaatt ttacacattg ccactaaacg tctaaaccct tgtaatttgt 12720
 ttttgtttta ctatgtgtgt tatgtatttg atttgcgata aatttttata tttggtacta 12780
 25 aatttataac accttttatg ctaacgtttg ccaacactta gcaatttgca agttgattaa 12840
 ttgattctaa attatttttg tcttctaaat acatatacta atcaactgga aatgtaaata 12900
 tttgctaata tttctactat aggagaatta aagtgagtga atatggtacc acaaggtttg 12960
 30 gagatttaat tgttgcaatg ctgcatggat ggcataata ccaaaccattc aataattctt 13020
 gaggataata atggtaccac acaagatttg aggtgcatga acgtcacgtg gacaaaagggt 13080
 35 ttagtaattt ttcaagacaa caatgttacc acacacaagt tttgagggtgc atgcatggat 13140
 gccctgtgga aagtttaaaa atattttgga aatgatttgc atggaagcca tgtgtaaaac 13200
 catgacatcc acttgaggga tgcaataatg aagaaaacta caaatttaca tgcaactagt 13260
 40 tatgcatgta gtctatataa tgaggatttt gcaatacttt cattcatata cactcactaa 13320
 gttttacacg attataattt cttcatagcc agcagatctg ccggcatcga tcccgggcca 13380
 45 tggcctgctt taatgagata tgcgagacgc ctatgatcgc atgatatttg ctttcaattc 13440
 tgttgtgcac gttgtaaaaa acctgagcat gtgtagctca gatccttacc gccggtttcg 13500
 gttcattcta atgaatatat capccgttac tatcgtattt ttatgaataa tattctccgt 13560
 50 tcaatttact gattgtccgt cgacgagctc ggcgcgcctc tagaggatcg atgaattcag 13620
 atcggctgag tggctccttc aacgttgccg ttctgtcagt tccaaacgta aaacggcttg 13680
 55 tcccgcgtca tcggcggggg tcataacgtg actcccttaa ttctccgctc atgatcagat 13740
 tgctgtttcc cgccttcagt ttaaaactatc agtgtttgac aggatatatt ggcgggtaaa 13800
 cctaagagaa aagagcgttt attagaataa tcggatattt aaaagggcgt gaaaagggtt 13860
 60 atccttcgtc catttgtatg tgcattgcaa ccacagggtt cccca 13905

```
<210> 32
<211> 1443
<212> DNA
<213> Phaeodactylum tricornutum
```

0. $\begin{matrix} <220>^+ \\ <221> \text{ CDS} \\ <222> (9) \dots (1442) \\ <223> \end{matrix}$

[illegible]

	Leu	Phe	Trp	Gly	Asn	Leu	Met	Gln	Gly	Tyr	Ser	Val	Gln	Trp	Trp	Lys	
					195					200						205	
5	aac	aag	cac	aac	gga	cac	cac	gcc	gtc	ccc	aac	ctc	cac	tgc	tcc	tcc	674
	Asn	Lys	His	Asn	Gly	His	His	Ala	Val	Pro	Asn	Leu	His	Cys	Ser	Ser	
				210					215					220			
10	gca	gtc	gcg	caa	gat	ggg	gac	ccg	gac	atc	gat	acc	atg	ccc	ctt	ctc	722
	Ala	Val	Ala	Gln	Asp	Gly	Asp	Pro	Asp	Ile	Asp	Thr	Met	Pro	Leu	Leu	
				225				230					235				
15	gcc	tgg	tcc	gtc	cag	caa	gcc	cag	tct	tac	cgg	gaa	ctc	caa	gcc	gac	770
	Ala	Trp	Ser	Val	Gln	Gln	Ala	Gln	Ser	Tyr	Arg	Glu	Leu	Gln	Ala	Asp	
				240			245					250					
20	gga	aag	gat	tcg	ggt	ttg	gtc	aag	ttc	atg	atc	cgt	aac	caa	tcc	tac	818
	Gly	Lys	Asp	Ser	Gly	Leu	Val	Lys	Phe	Met	Ile	Arg	Asn	Gln	Ser	Tyr	
						260					265					270	
25	ttt	tac	ttt	ccc	atc	ttg	ttg	ctc	gcc	cgc	ctg	tcg	tgg	ttg	aac	gag	866
	Phe	Tyr	Phe	Pro	Ile	Leu	Leu	Leu	Ala	Arg	Leu	Ser	Trp	Leu	Asn	Glu	
						275				280					285		
30	tcc	ttc	aag	tgc	gcc	ttt	ggg	ctt	gga	gct	gcg	tcg	gag	aac	gct	gct	914
	Ser	Phe	Lys	Cys	Ala	Phe	Gly	Leu	Gly	Ala	Ala	Ser	Glu	Asn	Ala	Ala	
					290				295					300			
35	ctc	gaa	ctc	aag	gcc	aag	ggt	ctt	cag	tac	ccc	ctt	ttg	gaa	aag	gct	962
	Leu	Glu	Leu	Lys	Ala	Lys	Gly	Leu	Gln	Tyr	Pro	Leu	Leu	Glu	Lys	Ala	
				305				310					315				
40	ggc	atc	ctg	ctg	cac	tac	gct	tgg	atg	ctt	aca	gtt	tcg	tcc	ggc	ttt	1010
	Gly	Ile	Leu	Leu	His	Tyr	Ala	Trp	Met	Leu	Thr	Val	Ser	Ser	Gly	Phe	
				320			325					330					
45	gga	cgc	ttc	tcg	ttc	gcg	tac	acc	gca	ttt	tac	ttt	cta	acc	gcg	acc	1058
	Gly	Arg	Phe	Ser	Phe	Ala	Tyr	Thr	Ala	Phe	Tyr	Phe	Leu	Thr	Ala	Thr	
						340					345				350		
50	gcg	tcc	tgt	gga	ttc	ttg	ctc	gcc	att	gtc	ttt	ggc	ctc	ggc	cac	aac	1106
	Ala	Ser	Cys	Gly	Phe	Leu	Leu	Ala	Ile	Val	Phe	Gly	Leu	Gly	His	Asn	
						355				360					365		
55	ggc	atg	gcc	acc	tac	aat	gcc	gac	gcc	cgt	ccg	gac	ttc	tgg	aag	ctc	1154
	Gly	Met	Ala	Thr	Tyr	Asn	Ala	Asp	Ala	Arg	Pro	Asp	Phe	Trp	Lys	Leu	
					370			375						380			
60	caa	gtc	acc	acg	act	cgc	aac	gtc	acg	ggc	gga	cac	ggt	ttc	ccc	caa	1202
	Gln	Val	Thr	Thr	Thr	Arg	Asn	Val	Thr	Gly	Gly	His	Gly	Phe	Pro	Gln	
				385				390					395				
65	gcc	ttt	gtc	gac	tgg	ttc	tgt	ggt	ggc	ctc	cag	tac	caa	gtc	gac	cac	1250
	Ala	Phe	Val	Asp	Trp	Phe	Cys	Gly	Gly	Leu	Gln	Tyr	Gln	Val	Asp	His	
				400			405				410						
70	cac	tta	ttc	ccc	agc	ctg	ccc	cga	cac	aat	ctg	gcc	aag	aca	cac	gca	1298
	His	Leu	Phe	Pro	Ser	Leu	Pro	Arg	His	Asn	Leu	Ala	Lys	Thr	His	Ala	
				415			420				425					430	
75	ctg	gtc	gaa	tcg	ttc	tgc	aag	gag	tgg	ggt	gtc	cag	tac	cac	gaa	gcc	1346
	Leu	Val	Glu	Ser	Phe	Cys	Lys	Glu	Trp	Gly	Val	Gln	Tyr	His	Glu	Ala	
					435				440						445		

gac ctt gtg gac ggg acc atg gaa gtc ttg cac cat ttg ggc agc gtg 1394
 Asp Leu Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val
 450 455 460
 5 gcc ggc gaa ttc gtc gtg gat ttt gta cgc gat gga ccc gcc atg taa a 1443
 Ala Gly Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
 465 470 475
 10 <210> 33
 <211> 477
 <212> PRT
 <213> *Phaeodactylum tricornutum*
 15 <400> 33
 Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala
 1 5 10 15
 20 Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp
 20 25 30
 25 Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His
 35 40 45
 30 Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met
 50 55 60
 35 Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met
 65 70 75 80
 40 Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
 85 90 95
 Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys
 100 105 110
 45 Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr
 115 120 125
 50 Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val
 130 135 140
 55 Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu
 145 150 155 160
 60 Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His
 165 170 175
 His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe

	180	185	190
5	Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys 195 200 205		
10	His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val 210 215 220		
15	Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp 225 230 235 240		
20	Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys 245 250 255		
25	Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr 260 265 270		
30	Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe 275 280 285		
35	Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu 290 295 300		
40	Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile 305 310 315 320		
45	Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg 325 330 335		
50	Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser 340 345 350		
55	Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met 355 360 365		
60	Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val 370 375 380		
	Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe 385 390 395 400		
	Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu 405 410 415		
	Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val 420 425 430		

Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu
 435 440 445

5

Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly
 450 455 460

10

Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
 465 470 475

15

<210> 34
 <211> 17061
 <212> DNA
 <213> *Phaeodactylum tricornutum*, *Physcomitrella patens*, *Caenorhabditis elegans*

20

<220>
 <221> CDS
 <222> (4554)..(5987)
 <223>

25

<220>
 <221> CDS
 <222> (2805)..(3653)
 <223>

30

<220>
 <221> CDS
 <222> (1026)..(1898)
 <223>

35

<400> 34
 tgggggaaccc tgtggttggc atgcacatac aaatggacga aggataaacc ttttcacgcc 60
 cttttaaata tccgattatt ctaataaacg ctcttttctc ttaggtttac ccgccaatat 120
 atcctgtcaa aactgatag tttaaactga aggcgggaaa cgacaatctg atcatgagcg 180
 45 gagaattaag ggagtcacgt tatgaccccc gccgatgacg cgggacaagc cgtttttacgt 240
 ttggaactga cagaaccgca acgttgaagg agccactcag ccgatctgaa ttcacgatc 300
 ctctagaggc gcgccgagct cctcgagcaa atttacacat tgccactaaa cgtctaaacc 360
 50 cttgtaattt gtttttgttt tactatgtgt gttatgtatt tgatttgca taaattttta 420
 tatttggtac taaatttata acacctttta tgctaacgtt tgccaacact tagcaatttg 480
 55 caagttgatt aattgattct aaattatttt tgtcttctaa atacatatac taatcaactg 540
 gaaatgtaaa tatttgctaa tatttctact ataggagaat taaagtgagt gaatatggta 600
 ccacaagggtt tggagattta attgttgcaa tgctgcatgg atggcatata caccaaacad 660
 60 tcaataattc ttgaggataa taatggtaac acacaagatt tgaggtgcat gaacgtcacg 720

	tggacaaaag gtttagtaat ttttcaagac aacaatgtta ccacacacaa gttttgaggt	780
	gcatgcatgg atgccctgtg gaaagtttaa aaatattttg gaaatgattt gcatggaagc	840
5	catgtgtaaa accatgacat ccacttggag gatgcaataa tgaagaaaac tacaaattta	900
	catgcāacta gttatgcatg tagtctatat aatgaggatt ttgcaatact ttcattcata	960
10	cacactcact aagttttaca cgattataat ttcttcatag ccagcccacc gcggtgggcg	1020
	gccgc atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc	1070
	Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val	
	1 5 10 15	
15	tcg cag ggc gtg aat gca ttg ctg ggt agt ttt ggg gtg gag ttg acg	1118
	Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr	
	20 25 30	
20	gat acg ccc act acc aaa ggc ttg ccc ctc gtt gac agt ccc aca ccc	1166
	Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro	
	35 40 45	
25	atc gtc ctc ggt gtt tct gta tac ttg act att gtc att gga ggg ctt	1214
	Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu	
	50 55 60	
30	ttg tgg ata aag gcc agg gat ctg aaa ccg cgc gcc tcg gag cca ttt	1262
	Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe	
	65 70 75	
35	ttg ctc caa gct ttg gtg ctt gtg cac aac ctg ttc tgt ttt gcg ctc	1310
	Leu Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu	
	80 85 90 95	
40	agt ctg tat atg tgc gtg ggc atc gct tat cag gct att acc tgg cgg	1358
	Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg	
	100 105 110	
45	tac tct ctc tgg ggc aat gca tac aat cct aaa cat aaa gag atg gcg	1406
	Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala	
	115 120 125	
50	att ctg gta tac ttg ttc tac atg tct aag tac gtg gaa ttc atg gat	1454
	Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp	
	130 135 140	
55	acc gtt atc atg ata ctg aag cgc agc acc agg caa ata agc ttc ctc	1502
	Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu	
	145 150 155	
60	cac gtt tat cat cat tct tca att tcc ctc att tgg tgg gct att gct	1550
	His Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala	
	160 165 170 175	
65	cat cac gct cct ggc ggt gaa gca tat tgg tct gcg gct ctg aac tca	1598
	His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser	
	180 185 190	
70	gga gtg cat gtt ctc atg tat gcg tat tac ttc ttg gct gcc tgc ctt	1646
	Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu	
	195 200 205	

	cga agt agc cca aag tta aaa aat aag tac ctt ttt tgg ggc agg tac	1694
	Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr	
	210 215 220	
5	ttg aca caa ttc caa atg ttc cag ttt atg ctg aac tta gtg cag gct	1742
	Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala	
	225 230 235	
10	tac tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag	1790
	Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys	
	240 245 250 255	
15	att ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt	1838
	Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe	
	260 265 270	
20	tac gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct	1886
	Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala	
	275 280 285	
	aaa act gag tga tctagaaggc ctcctgcttt aatgagatat gcgagacgcc	1938
	Lys Thr Glu	
	290	
25	tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa cctgagcatg	1998
	tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc acccgttact	2058
30	atcgtatttt tatgaataat attctccgtt caatttactg attgtccgtc gagcaaat	2118
	acacattgcc actaaacgtc taaacccttg taatttgttt ttgttttact atgtgtgtta	2178
	tgtatttgat ttgcgataaa tttttatatt tgggtactaaa tttataacac cttttatgct	2238
35	aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat tatttttgtc	2298
	ttctaaatac atatactaata caactggaaa tgtaaataatt tgctaataatt tctactatag	2358
40	gagaattaaa gtgagtgaat atggtaccac aaggtttgga gatttaattg ttgcaatgct	2418
	gcatggatgg catatacacc aaacattcaa taattcttga ggataataat ggtaccacac	2478
	aagatttgag gtgcatgaac gtcacgtgga caaaagggtt agtaattttt caagacaaca	2538
45	atgttaccac acacaagttt tgaggtgcat gcatggatgc cctgtggaaa gtttaaaaat	2598
	attttggaaa tgatttgcac ggaagccatg tgtaaaacca tgacatccac ttggaggatg	2658
50	caataatgaa gaaaactaca aatttacatg caactagtta tgcattgtat ctatataatg	2718
	aggattttgc aatactttca ttcatacaca ctactaagt tttacacgat tataatttct	2778
	tcatagccag cggatccgcc cacata atg gag aac ttc tgg tct att gtt gtg	2831
55	Met Glu Asn Phe Trp Ser Ile Val Val	
	295	
60	ttt ttt cta ctc tca att ctc ttc att tta tat aac ata tcg aca gta	2879
	Phe Phe Leu Leu Ser Ile Leu Phe Ile Leu Tyr Asn Ile Ser Thr Val	
	300 305 310 315	
	tgc cac tac tat atg cgg att tcg ttt tat tac ttc aca att tta ttg	2927
	Cys His Tyr Tyr Met Arg Ile Ser Phe Tyr Tyr Phe Thr Ile Leu Leu	

	320	325	330	
5	cat gga atg gaa gtt tgt gtt aca atg atc cct tct tgg cta aat ggg His Gly Met Glu Val Cys Val Thr Met Ile Pro Ser Trp Leu Asn Gly 335 340 345	2975		
10	aag ggt gct gat tac gtg ttt cac tcg ttt ttc tat tgg tgt aaa tgg Lys Gly Ala Asp Tyr Val Phe His Ser Phe Phe Tyr Trp Cys Lys Trp 350 355 360	3023		
15	act ggt gtt cat aca aca gtc tat gga tat gaa aaa aca caa gtt gaa Thr Gly Val His Thr Thr Val Tyr Gly Tyr Glu Lys Thr Gln Val Glu 365 370 375	3071		
20	ggt ccg gct gta gtt att tgt aat cat cag agt tct ctc gac att cta Gly Pro Ala Val Val Ile Cys Asn His Gln Ser Ser Leu Asp Ile Leu 380 385 390 395	3119		
25	tcg atg gca tca atc tgg ccg aag aat tgt gtt gta atg atg aaa cga Ser Met Ala Ser Ile Trp Pro Lys Asn Cys Val Val Met Met Lys Arg 400 405 410	3167		
30	att ctt gcc tat gtt cca ttc ttc aat ctc gga gcc tac ttt tcc aac Ile Leu Ala Tyr Val Pro Phe Phe Asn Leu Gly Ala Tyr Phe Ser Asn 415 420 425	3215		
35	aca atc ttc atc gat cga tat aac cgt gaa cgt gcg atg gct tca gtt Thr Ile Phe Ile Asp Arg Tyr Asn Arg Glu Arg Ala Met Ala Ser Val 430 435 440	3263		
40	gat tat tgt gca tct gaa atg aag aac aga aat ctt aaa ctt tgg gta Asp Tyr Cys Ala Ser Glu Met Lys Asn Arg Asn Leu Lys Leu Trp Val 445 450 455	3311		
45	ttt ccg gaa gga aca aga aat cgt gaa gga ggg ttc att cca ttc aag Phe Pro Glu Gly Thr Arg Asn Arg Glu Gly Phe Ile Pro Phe Lys 460 465 470 475	3359		
50	aaa gga gca ttc aat att gca gtt cgt gcg cag att ccc att att cca Lys Gly Ala Phe Asn Ile Ala Val Arg Ala Gln Ile Pro Ile Ile Pro 480 485 490	3407		
55	gtt gta ttc tca gac tat cgg gat ttc tac tca aag cca ggc cga tat Val Val Phe Ser Asp Tyr Arg Asp Phe Tyr Ser Lys Pro Gly Arg Tyr 495 500 505	3455		
60	ttc aag aat gat gga gaa gtt gtt att cga gtt ctg gat gcg att cca Phe Lys Asn Asp Gly Glu Val Val Ile Arg Val Leu Asp Ala Ile Pro 510 515 520	3503		
65	aca aaa ggg ctc act ctt gat gac gtc agc gag ttg tct gat atg tgt Thr Lys Gly Leu Thr Leu Asp Asp Val Ser Glu Leu Ser Asp Met Cys 525 530 535	3551		
70	cgg gac gtt atg ttg gca gcc tat aag gaa gtt act cta gaa gct cag Arg Asp Val Met Leu Ala Ala Tyr Lys Glu Val Thr Leu Glu Ala Gln 540 545 550 555	3599		
75	caa cga aat gcg aca cgg cgt gga gaa aca aaa gac ggg aag aaa tct Gln Arg Asn Ala Thr Arg Arg Gly Glu Thr Lys Asp Gly Lys Lys Ser 560 565 570	3647		

	gag taa gctagcgtta accctgcttt aatgagatat gcgagacgcc tatgatcgca	3703
	Glu	
5	tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa cctgagcatg tgtagctcag	3763
	atccttaccg ccggtttcgg ttcattctaa tgaatatatc acccgttact atcgtatttt	3823
10	tatgaataat attctccgtt caattttactg attgtccgtc gagcaaattt acacattgcc	3883
	actaaacgtc taaacccttg taattttgtt ttgttttact atgtgtgtta tgtatttgat	3943
	ttgcgataaa tttttatatt tgggtactaaa ttataacac cttttatgct aacgtttgcc	4003
15	aacacttagc aatttgcaag ttgattaatt gattctaaat tatttttgtc ttctaaatac	4063
	atatactaata caactggaaa tgtaaatatt tgctaataat tctactatag gagaattaaa	4123
20	gtgagtgaat atgggtaccac aagggttgga gatttaattg ttgcaatgct gcatggatgg	4183
	catatacacc aaacattcaa taattcttga ggataataat ggtaccacac aagatttgag	4243
	gtgcatgaac gtcacgtgga caaaagggtt agtaattttt caagacaaca atgttaccac	4303
25	acacaagttt tgagggtgcat gcatggatgc cctgtggaaa gtttaaaaat attttgaaa	4363
	tgatttgcac ggaagccatg tgtaaaacca tgacatccac ttggaggatg caataatgaa	4423
30	gaaaactaca aatttacatg caactagtta tgcatttagt ctatataatg aggattttgc	4483
	aatactttca ttcatacaca ctactaagt ttacacgat tataatttct tcatagccag	4543
	cagatctaaa atg ggc aaa gga ggg gac gct cgg gcc tcg aag ggc tca	4592
35	Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser	
	575 580 585	
	acg gcg gct cgc aag atc agt tgg cag gaa gtc aag acc cac gcg tct	4640
40	Thr Ala Ala Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser	
	590 595 600	
	ccg gag gac gcc tgg atc att cac tcc aat aag gtc tac gac gtg tcc	4688
	Pro Glu Asp Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser	
	605 610 615	
45	aac tgg cac gaa cat ccc gga ggc gcc gtc att ttc acg cac gcc ggt	4736
	Asn Trp His Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly	
	620 625 630	
50	gac gac atg acg gac att ttc gct gcc ttt cac gca ccc gga tcg cag	4784
	Asp Asp Met Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln	
	635 640 645	
	tcg ctc atg aag aag ttc tac att ggc gaa ttg ctc ccg gaa acc acc	4832
55	Ser Leu Met Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr	
	650 655 660 665	
	ggc aag gag ccg cag caa atc gcc ttt gaa aag ggc tac cgc gat ctg	4880
	Gly Lys Glu Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu	
	670 675 680	
60	cgc tcc aaa ctc atc atg atg ggc atg ttc aag tcc aac aag tgg ttc	4928
	Arg Ser Lys Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe	

	685	690	695	
5	tac gtc tac aag tgc ctc agc aac atg gcc att tgg gcc gcc gcc tgt Tyr Val Tyr Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys 700 705 710			4976
10	gct ctc gtc ttt tac tcg gac cgc ttc tgg gta cac ctg gcc agc gcc Ala Leu Val Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala 715 720 725			5024
15	gtc atg ctg gga aca ttc ttt cag cag tcg gga tgg ttg gca cac gac Val Met Leu Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp 730 735 740 745			5072
20	ttt ctg cac cac cag gtc ttc acc aag cgc aag cac ggg gat ctc gga Phe Leu His His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly 750 755 760			5120
25	gga ctc ttt tgg ggg aac ctc atg cag ggt tac tcc gta cag tgg tgg Gly Leu Phe Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp 765 770 775			5168
30	aaa aac aag cac aac gga cac cac gcc gtc ccc aac ctc cac tgc tcc Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser 780 785 790			5216
35	tcc gca gtc gcg caa gat ggg gac ccg gac atc gat acc atg ccc ctt Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu 795 800 805			5264
40	ctc gcc tgg tcc gtc cag caa gcc cag tct tac cgg gaa ctc caa gcc Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala 810 815 820 825			5312
45	gac gga aag gat tcg ggt ttg gtc aag ttc atg atc cgt aac caa tcc Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser 830 835 840			5360
50	tac ttt tac ttt ccc atc ttg ttg ctc gcc cgc ctg tcg tgg ttg aac Tyr Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn 845 850 855			5408
55	gag tcc ttc aag tgc gcc ttt ggg ctt gga gct gcg tcg gag aac gct Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala 860 865 870			5456
60	gct ctc gaa ctc aag gcc aag ggt ctt cag tac ccc ctt ttg gaa aag Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys 875 880 885			5504
65	gct ggc atc ctg ctg cac tac gct tgg atg ctt aca gtt tcg tcc ggc Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly 890 895 900 905			5552
70	ttt gga cgc ttc tcg ttc gcg tac acc gca ttt tac ttt cta acc gcg Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala 910 915 920			5600
75	acc gcg tcc tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc ggc cac Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His 925 930 935			5648

	aac ggc atg gcc acc tac aat gcc gac gcc cgt ccg gac ttc tgg aag Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys 940 945 950	5696
5	ctc caa gtc acc acg act cgc aac gtc acg ggc gga cac ggt ttc ccc Leu Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro 955 960 965	5744
10	caa gcc ttt gtc gac tgg ttc tgt ggt ggc ctc cag tac caa gtc gac Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp 970 975 980 985	5792
15	cac cac tta ttc ccc agc ctg ccc cga cac aat ctg gcc aag aca cac His His Leu Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His 990 995 1000	5840
20	gca ctg gtc gaa tgc ttc tgc aag gag tgg ggt gtc cag tac cac Ala Leu Val Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His 1005 1010 1015	5885
	gaa gcc gac ctt gtg gac ggg acc atg gaa gtc ttg cac cat ttg Glu Ala Asp Leu Val Asp Gly Thr Met Glu Val Leu His His Leu 1020 1025 1030	5930
25	ggc agc gtg gcc ggc gaa ttc gtc gtg gat ttt gta cgc gat gga Gly Ser Val Ala Gly Glu Phe Val Val Asp Phe Val Arg Asp Gly 1035 1040 1045	5975
30	ccc gcc atg taa agatctgccg gcatcgatcc cgggccatgg cctgctttaa Pro Ala Met	6027
	tgagatatgc gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacggt	6087
35	gtaaaaaacc tgagcatgtg tagctcagat ccttaccgcc ggtttcgggtt cattctaattg aatatatcac ccgttactat cgtatttttta tgaataatat tctccgttca atttactgat	6147 6207
40	tgtccgtcga cgagctcggc gcgccgtcga cctgcaggca tgcaagcttc acgctgccgc aagcactcag ggcgcaaggg ctgctaaagg aagcggaaca cgtagaaagc cagtccgcag	6267 6327
	aaacgggtgct gaccccgat gaatgtcagc tactgggcta tctggacaag ggaaaacgca	6387
45	agcgcaaaga gaaagcaggt agcttgacgt gggcttacat ggcatagct agactgggcg gttttatgga cagcaagcga accggaattg ccagctgggg cgccctctgg taagggtggg	6447 6507
50	aagccctgca aagtaaaactg gatggctttc ttgccgcaa ggatctgatg gcgcagggga tcaagatcat gagcggagaa ttaagggagt cacgttatga ccccgccga tgacgcggga	6567 6627
	caagccgttt tacgtttgga actgacagaa ccgcaacgtt gaaggagcca ctacgccgcg	6687
55	ggtttctgga gtttaattgag ctaagcacat acgtcagaaa ccattattgc gcgttcaaaa gtgcgctaag gtcactatca gctagcaaat atttctgtc aaaaatgtc cactgacgtt	6747 6807
60	ccataaatc cctcggtat ccaattagag tctcatattc actctcaatc cagatctcga ctctagtcga gggcccatgg gagcttggat tgaacaagat ggattgcacg caggttctcc	6867 6927

	ggccgcttgg	gtggagaggc	tattcggcta	tgactgggca	caacagacaa	tcggctgctc	6987
	tgatgccgcc	gtgttccggc	tgtcagcgca	ggggcgcccc	gttctttttg	tcaagaccga	7047
5	cctgtccggt	gccctgaatg	aactgcagga	cgaggcagcg	cggctatcgt	ggctggccac	7107
10	gacgggcggt	ccttgccgag	ctgtgctcga	cgttgctact	gaagcgggaa	gggactggct	7167
	gctattgggc	gaagtgccgg	ggcaggatct	cctgtcatct	caccttgctc	ctgccgagaa	7227
	agtatcoatc	atggctgatg	caatgcggcg	gctgcatacg	cttgatccgg	ctacctgccc	7287
	attcgaccac	caagcgaaac	atcgcatcga	gcgagcacgt	actcggatgg	aagccgggtct	7347
15	tgatgatcag	gatgatctgg	acgaagagca	tcaggggctc	gcgccagccg	aactgttcgc	7407
	caggctcaag	gcgcgcatgc	ccgacggcga	ggatctcgtc	gtgacccatg	gcgatgcctg	7467
20	cttgccgaat	atcatggtgg	aaaatggccg	cttttctgga	ttcatcgact	gtggccggct	7527
	gggtgtggcg	gaccgctatc	aggacatagc	gttggctacc	cgtgatattg	ctgaagagct	7587
	tggcggcgaa	tgggctgacc	gcttcctcgt	gctttacggt	atcgccgctc	ccgattcgca	7647
25	gcgcacgcgc	ttctatcgcc	ttcttgacga	gttcttctga	gcgggaccca	agctagcttc	7707
	gacggatccc	ccgatgagct	aagctagcta	tatcatcaat	ttatgtatta	cacataatat	7767
30	cgcactcagt	ctttcatcta	cggcaatgta	ccagctgata	taatcagtta	ttgaaatatt	7827
	tctgaattta	aacttgcatc	aataaattta	tgtttttgct	tggactataa	tacctgactt	7887
	gttattttat	caataaatat	ttaaactata	tttctttcaa	gatgggaatt	aattcactgg	7947
35	ccgtcgtttt	acaacgtcgt	gactgggaaa	accctggcgt	tacccaactt	aatcgccttg	8007
	cagcacatcc	ccctttcgcc	agctggcgta	atagcgaaga	ggcccgccacc	gatcgccctt	8067
40	cccaacagtt	gcgcagcctg	aatggcgccc	gctcctttcg	ctttcttccc	ttcctttctc	8127
	gccacgttcg	ccggctttcc	ccgtcaagct	ctaaatcggg	ggctcccttt	agggttccga	8187
	tttagtgctt	tacggcacct	cgaccccaaa	aaacttgatt	tgggtgatgg	ttcacgtagt	8247
45	gggccatcgc	cctgatagac	ggtttttgcg	cctttgacgt	tggagtccac	gttcttttaat	8307
	agtggactct	tgttccaaac	tgggaacaaca	ctcaacccta	tctcgggcta	ttcttttgat	8367
50	ttataaggga	ttttgccgat	ttcggaacca	ccatcaaaca	ggatttttcg	ctgctggggc	8427
	aaaccagcgt	ggaccgcttg	ctgcaactct	ctcagggcca	ggcgggtgaag	ggcaatcagc	8487
	tgttgcccgt	ctcactgggt	aaaagaaaaa	ccaccccagt	acattaaaaa	cgtccgcaat	8547
55	gtgttattaa	gttgtctaag	cgtcaatttg	tttacaccac	aatatatact	gccaccagcc	8607
	agccaacagc	tccccgaccg	gcagctcggc	acaaaatcac	cactcgatac	aggcagccca	8667
60	tcagtccggg	acggcgctcag	cgggagagcc	gttgtaaggc	ggcagacttt	gctcatgtta	8727
	ccgatgctat	tcggaagaac	ggcaactaag	ctgccgggtt	tgaaacacgg	atgatctcgc	8787

ggagggtagc atgttgattg taacgatgac agagcgttgc tgcctgtgat caaatatcat 8847
 ctccctcgca gagatccgaa ttatcagcct tcttattcat ttctcgctta accgtgacag 8907
 5 gctgtcgatc ttgagaacta tgccgacata ataggaaatc gctggataaa gccgctgagg 8967
 aagctgagtg gcgctatttc tttagaagtg aacgttgacg atatcaactc ccctatccat 9027
 10 tgctcaccga atggtacagg tcggggaccc gaagttccga ctgtcggcct gatgcatccc 9087
 cggctgatcg accccagatc tggggctgag aaagcccagt aaggaaacaa ctgtaggttc 9147
 gagtcgcgag atcccccgga accaaaggaa gtaggttaaa cccgctccga tcaggccgag 9207
 15 ccacgccagg ccgagaacat tggttcctgt aggcacggg attggcggat caaacactaa 9267
 agctactgga acgagcagaa gtcctccggc cgccagttgc caggcggtaa aggtgagcag 9327
 20 aggcacggga ggttgccact tgccgggtcag caggtttccg aacgccatgg aaaccgcccc 9387
 cgccaggccc gctgcgacgc cgacaggatc tagcgtgctg tttggtgtca acaccaacag 9447
 cgccaqgccc gcagttccgc aaatagcccc caggaccgcc atcaatcgta tcgggctacc 9507
 25 tagcagagcg gcagagatga acacgaccat cagcggctgc acagcgcta ccgtcgccgc 9567
 gaccccgccc ggcaggcggg agaccgaaat aaacaacaag ctccagaata gcgaaatatt 9627
 aagtgcgccg aggatgaaga tgcgcatcca ccagattccc gttggaatct gtcggacgat 9687
 30 catcacgagc aataaaccgc ccggcaacgc ccgcagcagc ataccggcga cccctcggcc 9747
 tcgctgttcg ggctccacga aaacgccgga cagatgcgcc ttgtgagcgt ccttggggcc 9807
 35 gtcctcctgt ttgaagaccg acagcccaat gatctcgccg tcgatgtagg cgccgaatgc 9867
 cacggcatct cgcaaccgtt cagcgaacgc ctccatgggc tttttctcct cgtgctcgta 9927
 40 aacggacccg aacatctctg gagctttctt cagggccgac aatcggatct cgcggaaatc 9987
 ctgcacgtcg gccgctcaa gccgtcgaat ctgagcctta atcacaattg tcaattttaa 10047
 tctctgttt atcggcagtt cgtagagcgc gccgtgcgtc ccgagcgata ctgagcgaag 10107
 45 caagtgcgtc gagcagtgcc cgcttgttcc tgaaatgcca gtaaagcgct ggctgctgaa 10167
 ccccgagccg gaactgacc cacaaggccc tagcgtttgc aatgcaccag gtcattcattg 10227
 50 acccaggcgt gttccaccag gccgctgcct cgcaactctt cgcaggcttc gccgacctgc 10287
 tcgcgccact tcttcacgcg ggtggaatcc gatccgcaca tgaggcggaa ggtttccagc 10347
 ttgagcgggt acggctcccg gtgcgagctg aaatagtcga acatccgtcg ggccgtcggc 10407
 55 gacagcttgc ggtaactctc ccatatgaat ttcgtgtagt ggtcgccagc aaacagcacg 10467
 acgatttcct cgtcgatcag gacctggcaa cgggacgttt tcttgccacg gtccaggacg 10527
 60 cggaagcggg gcagcagcga caccgattcc aggtgccccaa cgcggtcggg cgtgaagccc 10587
 atcgccgtcg cctgtaggcg cgacaggcat tctcggcct tcgtgtaata ccggccattg 10647

atcgaccagc ccaggctcctg gcaaagctcg tagaacgtga aggtgatcgg ctcgccgata 10707
 ggggtgcgct tcgctactc caacacctgc tgccacacca gtctgtcatc gtcggcccg 10767
 5 agctcgacgc cgggtgtaggt gatcttcacg tccttggtga cgtggaaaat gacctgttt 10827
 tgcagcgcct cgcgcgggat tttcttggtg cgcgtggtga acagggcaga gcgggccgtg 10887
 10 tcgtttggca tcgctcgcat cgtgtccggc cagggcgcaa tatcgaacaa ggaaagctgc 10947
 atttcttga tctgtctgtt cgtgtgtttc agcaacgcgg cctgcttggc ctgctgacc 11007
 tgttttgcca ggtcctcgcc ggcggttttt cgcttcttgg tcgtcatagt tctcgcgtg 11067
 15 tcgatggtca tcgacttcgc caaacctgcc gcctcctgtt cgagacgacg cgaacgctcc 11127
 acggcggccg atggcgcggg cagggcaggg ggagccagtt gcacgctgtc gcgctcgatc 11187
 20 ttggccgtag cttgctggac catcgagccg acggactgga aggtttcgcg ggcgcacgc 11247
 atgacggtgc ggcttgcat ggtttcggca tcctcggcgg aaaacccgc gtcgatcagt 11307
 tcttgdtgt atgccttcg gtcaaacgtc cgattcattc accctcctt cgggattgcc 11367
 25 ccgactcacg ccggggcaat gtgcccttat tcctgatttg accgccttg tgcttggtg 11427
 tccagataat ccaccttatc ggcaatgaag tcggtccgt agaccgtctg gccgtccttc 11487
 30 tcgtacttgg tattccgaat cttgccctgc acgaatacca gcgaccctt gcccaaatac 11547
 ttgccgtggg cctcggcctg agagccaaaa cacttgatgc ggaagaagtc ggtgcgctcc 11607
 tgcttgctgc cggcatcgtt gcgccacatc taggtactaa aacaattcat ccagtaaaat 11667
 35 ataataatattt atttctccc aatcaggctt gatccccagt aagtcaaaaa atagctcgac 11727
 atactgttct tcccgatat cctccctgat cgaccggacg cagaaggcaa tgtcatacca 11787
 40 cttgtccgcc ctgcgcttc tccaagatc aataaagcca cttactttgc catctttcac 11847
 aaagatgttg ctgtctccca ggtcgcctg ggaaaagaca agttcctctt cgggcttttc 11907
 cgtctttaa aaatcataca gctcgcgcgg atctttaa at ggagtgtctt ctccagtt 11967
 45 ttcgcaatcc acatcgcca gatcgttatt cagtaagtaa tccaattcgg ctaagcggct 12027
 gtctaagcta ttcgtatagg gacaatccga tatgtcgatg gagtgaaaga gcctgatgca 12087
 50 ctccgcatag agctcgataa tcttttcagg gctttgttca tcttcatact ctccgagca 12147
 aaggacgcca tcggcctcac tcatgagcag attgctccag ccatcatgcc gttcaaagt 12207
 caggacctt ggaacaggca gctttccttc cagccatagc atcatgtcct tttccgcttc 12267
 55 cacatcatag gtggtccctt tataccggct gtccgtcatt tttaaataa ggttttcatt 12327
 ttctcccacc agcttatata ccttagcagg agacattcct tcggtatctt ttacgcagcg 12387
 60 gtatttttcg atcagttttt tcaattccgg tgatattctc attttagcca tttattattt 12447
 ccttcctctt ttctacagta tttaaagata cccaagaag ctaattataa caagacgaac 12507

tccaattcac tgttccttgc attctaaaac cttaaatacc agaaaacagc tttttcaaag 12567
ttgttttcaa agttggcgta taacatagta tcgacggagc cgattttgaa accacaatta 12627
5 tgggtgatgc tgccaactta ctgatttagt gtatgatggt gtttttgagg tgctccagt 12687
gcttctgtgt ctatcagctg tccctcctgt tcagctactg acgggggtggt gcgtaacggc 12747
10 aaaagcaccg ccggacatca gcgctatctc tgctctcact gccgtaaaac atggcaactg 12807
cagttcactt acaccgcttc tcaaccdggg acgcaccaga aaatcattga tatggccatg 12867
aatggcggtg gatgccgggc aacagcccgc attatgggcg ttggcctcaa cacgatttta 12927
15 cgtcacttaa aaaactcagg ccgcagtcgg taacctcgcg catacagccg ggcagtgcg 12987
tcctcgtctg cgcggaaatg gacgaacagt ggggctatgt cggggctaaa tcgcgccagc 13047
gctggctggt ttacgcgtat gacagtctcc ggaagacggt tgttgccgac gtattcggtg 13107
20 aacgcactat ggcgacgctg gggcgctctta tgagcctgct gtcacccttt gacgtggtga 13167
tatggatgac ggatggctgg ccgctgtatg aatcccgct gaagggaag ctgcacgtaa 13227
25 tcagcaagcg atatacgag cgaattgagc ggcataacct gaatctgagg cagcacctgg 13287
cacggctggg acggaagtcg ctgtcgttct caaaatcggg ggagctgcat gacaaagtca 13347
tcgggcatta totgaacata aaacactatc aataagttgg agtcattacc caattatgat 13407
30 agaatttaca agctataagg ttattgtcct gggtttcaag cattagtcca tgcaagtttt 13467
tatgctttgc ccattctata gatatttga taagcgcgt gcctatgcct tgccccctga 13527
35 aatccttaca tacggcgata tcttctatat aaaagatata ttatcttacc agtattgtca 13587
atatattcaa ggcaatctgc ctctcatcc tcttcacct ctctgtcttg gtagcttttt 13647
40 aaatatggcg cttcatagag taattctgta aaggccaat tctcgttttc atacctcggt 13707
ataatcttac ctatcacctc aaatggttcg ctgggtttat cgcacccccg aacacgagca 13767
cggcaccgcg gaccactatg ccaagaatgc ccaaggtaaa aattgccggc cccgccatga 13827
45 agtcggtgaa tgccccgacg gccgaagtga agggcaggcc gccaccagg ccgcccct 13887
cactgcccgg cacctggtcg ctgaatgtcg atgccagcac ctgcggcacg tcaatgcttc 13947
cgggcgtcgc gtcgggctg atcgcccatc ccgttactgc ccgatcccg gcaatggcaa 14007
50 ggactgccag cgctgccatt tttggggtga ggccgttcgc ggccgagggg cgcagcccct 14067
ggggggatgg gagggccgcg ttagcgggcc gggagggttc gagaaggggg ggcaccccc 14127
55 ttcggcgtgc gcggtcacgc gcacaggcg cagccctggt taaaaacaag gttataaat 14187
attgggttaa aagcagggtt aaagacaggt tagcgggtggc cgaaaaacgg gcggaaaccc 14247
60 ttgcaaatgc tggattttct gcctgtggac agcccccaa atgtcaatag gtgcgcccct 14307
catctgtcag cactctgcc ctcaagtgtc aaggatcgcg cccctcatct gtcagtagtc 14367

ggcggccctca agtgtcaata ccgcaggggca cttatcccca ggcttgtcca catcatctgt 14427
 gggaaactcg cgtaaaatca ggcgttttcg ccgatttgcg aggctggcca gctccacgtc 14487
 5 gcccggccgaa atcgagcctg cccctcatct gtcaacgccg cgcgggtga gtcggccct 14547
 caagtgtcaa cgtccgcccc tcatctgtca gtgagggcca agttttccgc gaggtatcca 14607
 10 caacgccggc ggccgcggtg tctgcacac ggcttcgacg gcgtttctgg cgcgtttgca 14667
 gggccataga cggccgccag ccagcggcg agggcaacca gcccggtgag cgtcgaaag 14727
 gcgctcggtc ttgccttgtc cgtcggtgat gtacttcacc agctccgca agtcgctctt 14787
 15 cttgatggag cgcagggga cgtgcttggc aatcacgcgc acccccggc cgttttagcg 14847
 gctaaaaaag tcatggctct gccctcgggc ggaccacgcc catcatgacc ttgccaagct 14907
 20 cgtcctgctt ctcttcgac ttcgccagca gggcgaggat cgtggcatca ccgaaccgcg 14967
 ccgtgcgagg gtcgtcggtg agccagagtt tcagcaggcc gccagggcg cccaggtcgc 15027
 cattgatgag ggccagctcg cggacgtgct catagtccac gacgcccgtg atttttagc 15087
 25 cctggccgac ggccagcagg taggcccaga ggctcatgcc ggccgcccgc gccttttct 15147
 caatcgctct tcgttcgtct ggaaggcagt acacctgat aggtgggctg cccttctg 15207
 30 ttggcttggg ttcacagcc atccgcttgc cctcatctgt tacgccggcg gtagccggcc 15267
 agcctcgag agcaggattc ccgttgagca ccgccagggt cgaataaggg acagtgaaga 15327
 aggaacaccc gtcgcgggt gggcctactt cacctatcct gcccggtga cgcggttga 15387
 35 tacaccaagg aaagtctaca cgaaccctt ggcaaaatcc tgtatatcgt gcgaaaagg 15447
 atggatatac cgaaaaatc gctataatga cccgaagca gggttatgca gcggaaaagc 15507
 40 gccacgcttc ccgaaggag aaaggcgag aggtatccgg taagcggcag ggtcggaaca 15567
 ggagagcgca cgaggagct tccaggggga aacgcctggt atctttatag tcctgtcggg 15627
 tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta 15687
 45 tggaaaaacg ccagcaacgc ggccttttta cgttcctggt ccttttgctg gccttttgc 15747
 cacatgttct ttctgcgtt atccctgat tctgtggata accgtattac cgcctttgag 15807
 50 tgagctgata ccgctcgcc cagccgaac accgagcgca gcgagtcagt gagcgaggaa 15867
 gcggaagagc gccagaagg cgcagagag gccgagcgcg gccgtgaggc ttggacgcta 15927
 gggcagggca tgaaaaagc cgtagcggc tgctacggc gtctgacgc gtggaaagg 15987
 55 ggaggggatg ttgtctacat ggctctgct tagtgagtgg gttgcgtcc ggcagcggtc 16047
 ctgatcaatc gtcaccctt ctcggtcctt caacgttct gacaacgagc ctcttttcg 16107
 60 ccaatccatc gacaatcacc gcgagtcct gctgaacgc tgcgtccgga ccggttcgt 16167
 cgaaggcgtc tatcgcgcc cgcaacagc gcgagagcg agcctgttca acggtgccg 16227

cgcgctcgcc ggcacgcgtg tcgcccggcct gctcctcaag cacggcccca acagtgaagt 16287
 agctgattgt catcagcgca ttgacggcgt ccccgccga aaaacccgcc tcgcagagga 16347
 5 agcgaagctg cgcgtcgcc gtttccatct gcggtgcgc cggtcgcgtg ccggcatgga 16407
 tgccgcgcgc atcgcggtag gcgagcagcg cctgcctgaa gctgcgggca ttcccgatca 16467
 10 gaaatgagcg ccagtcgctg tcggctctcg gcaccgaatg cgtatgattc tccgccagca 16527
 tggcttcggc cagtgcgtcg agcagcgccc gcttgctcct gaagtgccag taaagcgccg 16587
 gctgctgaac cccaaccgt tccgccagtt tgcgtgctgt cagaccgtct acgccgacct 16647
 15 cgttcaacag gtccagggcg gcacggatca ctgtattcgg ctgcaacttt gtcattgcttg 16707
 acactttatc actgataaac ataatatgtc caccaactta tcagtataa agaatccgcg 16767
 20 cgttcaatcg gaccagcgga ggctggtccg gaggccagac gtgaaacca acatacccct 16827
 gatcgtaatt ctgagcactg tcgcgctcga cgctgtcggc atcggcctga ttatgccggt 16887
 gctgccgggc ctctgcgcg atctggttca ctgcaacgac gtcaccgccc actatggcat 16947
 25 tctgctggcg ctgtatgcgt tgggtgaatt tgctgcgc cctgtgctgg gcgcgctgtc 17007
 ggatcgtttc gggcgccggc caatcttgct cgtctcgtg gccggcgcca gatc 17061

30 <210> 35
 <211> 290
 <212> PRT
 <213> *Phaeodactylum tricornutum*, *Physcomitrella patens*, *Caenorhabditis elegans*

35 <400> 35

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser
 1 5 10 15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp
 20 25 30

45 Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile
 35 40 45

50 Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu
 50 55 60

55 Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu
 65 70 75 80

60 Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser
 85 90 95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr

100 105 110

5 Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile
115 120 125

10 Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr
130 135 140

15 Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His
145 150 155 160

20 Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His
165 170 175

25 His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly
180 185 190

30 Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg
195 200 205

35 Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu
210 215 220

40 Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr
225 230 235 240

45 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile
245 250 255

50 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr
260 265 270

55 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys
275 280 285

Thr Glu
290

<210> 36
<211> 282
<212> PRT
<213> Phaeodactylum tricornutum, Physcomitrella patens, Caenorhabditis elegans
<400> 36

60 Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
1 5 10 15

Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
 20 25 30

5 Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
 35 40 45

10 Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
 50 55 60

15 His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
 65 70 75 80

Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
 85 90 95

20 Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110

25 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125

30 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140

35 Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160

Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
 165 170 175

40 Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
 180 185 190

45 Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
 195 200 205

50 Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
 210 215 220

55 Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
 225 230 235 240

Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255

60 Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg

260

265

270

5 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
275 280

<210> 37
<211> 477

10 <212> PRT

<213> *Phaeodactylum tricornutum*, *Physcomitrella patens*, *Caenorhabditis elegans*

15 <400> 37

Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala
1 5 10 15

20 Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp
20 25 30

25 Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His
35 40 45

30 Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met
50 55 60

Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met
65 70 75 80

35 Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu
85 90 95

40 Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys
100 105 110

45 Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr
115 120 125

50 Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val
130 135 140

Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu
145 150 155 160

55 Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His
165 170 175

60 His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe
180 185 190

Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys
 195 200 205
 5 His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val
 210 215 220
 10 Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp
 225 230 235 240
 15 Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys
 245 250 255
 20 Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr
 260 265 270
 Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe
 275 280 285
 25 Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu
 290 295 300
 30 Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile
 305 310 315 320
 35 Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg
 325 330 335
 40 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser
 340 345 350
 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met
 355 360 365
 45 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val
 370 375 380
 50 Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe
 385 390 395 400
 55 Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu
 405 410 415
 60 Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val
 420 425 430
 Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu

435

440

445

5 Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly
450 455 460

10 Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met
465 470 475